

الدكتور - محمد ماقول

الامضاء الوطني



ديوان المندوبيات العامة

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

الدكتور
محمد راتول

الإحصاء الوصفي

الطبعة الثانية



ديوان المطبوعات الجامعية

الساحة المركزية - بن عكنون - الجزائر
economicrg.blogspot.com
الباحث الاقتصادي

Economic **R**esearcher **G**ate

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

© ديوان المطبوعات الجامعية 2006-12

رقم التسجيل: 4.01.4624

رقم الترخيص: 9961.0. 0836.7 (ISBN)

رقم الإيداع القانوني: 2005/1242

economicrg.blogspot.com
groups/economicrg
الباحث الاقتصادي

بوابة
Economic Research Gate

بسم الله الرحمن الرحيم

الإهداء

الى الأبناء محمد وعبد المالك

مليكة وفاطمة الزهراء

الى أمهم، أجدادهم وجداتهم

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

فصول الكتاب

3	الفصل الأول: مفاهيم، استخدامات و منهجية
31	الفصل الثاني: تمييز البيانات الإحصائية
55	الفصل الثالث: العرض البياني
81	الفصل الرابع: مقاييس النزعة المركزية
137	الفصل الخامس: مقاييس التشتت
153	الفصل السادس: أشكال التوزيعات التكرارية
169	الفصل السابع: الإنحدار والارتباط
203	الفصل الثامن: السلاسل الزمنية
241	الفصل التاسع : الأرقام القياسية و معدلات النمو

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

مقدمة

الكتاب الذي بين يديك عبارة عن مجموعة من المواضيع المنقحة في الإحصاء الوصفي، قدمتها كمحاضرات في عدة معاهد وكليات لعدة تخصصات في العلوم الإنسانية و بخاصة تخصصات العلوم الاقتصادية و علوم التسيير و تخصصات المدى القصير في المحاسبة و الضرائب و الإعلام الآلي للتسيير، وهذا على فترات متلاحقة، وقد حرصت في تقديمها من خلال هذا الكتاب على البساطة في السرد والمنهجية في العرض، وعلى التوضيح عن طريق الأمثلة المباشرة في معظم المواضيع، حرصا على الإيضاح والفهم السريع، والإستيعاب الجيد، وذلك عبر تسعة فصول أساسية معروفة في تقديم مقياس الإحصاء الوصفي، مراعيًا الفترة الزمنية اللازمة لتدريس هذا المقياس للمستويات الجامعية، وهي إما سداسي واحد، بحجم ثلاث ساعات محاضرات و ساعتين أعمال موجهة أسبوعيا أو سنويا بحجم ساعة و نصف محاضرات و ساعة و نصف أعمال موجهة، و أنهيت كل فصل بمجموعة من التمارين بعض معطياتها إفتراضي و الآخر عبارة عن بعض المؤشرات الاقتصادية و الإجتماعية مصدرها الديوان الوطني للإحصائيات أساسا، تكون عبارة عن سلاسل من التمارين يلزم الطلبة على تحضيرها لتكون قاعة عمل في حصص الأعمال الموجهة، وقد تعتمد تقديم كل العلاقات الرياضية بالحروف اللاتينية، لتكون معظم المراجع المتواجدة في المكتبات تقدمها عادة بهذه الحروف إضافة الى أنها شائعة الإستخدام في معظم التخصصات العلمية الجامعية .

وإذ أقدم هذا العمل المتواضع لطلبتنا في الجذوع المشتركة في شتى العلوم الإنسانية و في بعض التخصصات من الطور الثانوي، فإنني أرجو أن أكون قد قدمت عملاً مفيداً في تقريب مادة الإحصاء الوصفي من الطالب، وتوفير مرجع إضافي في مكتبتنا، كما أرجو أن يكون في مستوى الجهد المبذول في تأليفه وإخراجه على هذا النحو.

و حيث أن بداية كل عمل قد تشوبها بعض النقائص، فإنني أنتظر من القاريء كل ملاحظة أو تنبيه على خطأ أو رأي في المنهجية أو غير ذلك، لنتمكن من تنقيحه وتحسينه وإستدراك نواقصه في طبعة موالية إن شاء الله.

الأستاذ : محمد راتول.

الفصل الأول

الإحصاء، مفاهيم، استخدامات و منهجية.

كلمة الإحصاء مصطلح شائع الاستخدام في مختلف مجالات العلوم، خاصة العلوم الاقتصادية والتجارية و علوم التسيير عامة، إذ أن معظم القرارات في هذا المجال، تعتمد على الدراسات الإحصائية، أو تستخدم الإحصاء في استدلالاتها، ويهدف هذا الفصل الى التعريف بالإحصاء، ومجالات استخدامه ومنهجية إجراء البحوث الإحصائية و مفرداتها.

أولاً : مفهوم الإحصاء : قبل الولوج في تقديم الإحصاء الوصفي كمادة تستند الى أسس علمية ثابتة ومحددة، لابد أولاً من توضيح المدلول العلمي للإحصاء بصفة عامة، و ما يتصل به، لكون كلمة الإحصاء كثيراً ما تتداخل مع كلمة تعداد أو كلمة إحصائيات، أو إحصاءات لدى الكثير من الطلبة، على الرغم من الفرق الشاسع بين هذه المصطلحات، وإزالة هذا التداخل نعطي التعاريف التالية:

1- تعداد : يقصد بكلمة التعداد في أدب المعاملات في البحوث الاقتصادية والاجتماعية و الإنسانية عامة، عملية العد التي تقوم بها أجهزة مختصة تابعة لهيئات رسمية (هيئات الدولة) في الغالب، و ذلك بهدف الحصول على معطيات حول ظاهرة أو مجموعة من الظواهر لغرض محدد أو غير محدد، فهو إذن عملية الحصر الكمي للظواهر.

2- إحصائيات : يقصد بكلمة إحصائيات (إحصاءات)، كل المعلومات العددية المتعلقة بظواهر أو نشاطات معينة، والتي قد تكون مقدمة في شكل جداول أو أشكال إحصائية مختلفة، سواء كان هذا التقديم منشوراً في كتيبات خاصة أو في مجلات أو دوريات، أو وثائق إدارية أو غيرها، وعلى هذا فإننا نعتبر

مثلا، أن الأرقام المقدمة في مجلة ما والمتعلقة بتطور السكان، أو المتعلقة بالإنتاج الداخلي الإجمالي، أو المتعلقة بمبيعات مؤسسة ما، إحصائيات، فالإحصائيات إذن هي نتيجة عملية العد.

تعريف 1-1 : يقصد بالإحصائيات كل المعطيات الرقمية، المتعلقة بظواهر إقتصادية أو إجتماعية... التي تجمعها هيئات مختصة سواء على مستوى النشاط أو على مستوى مجموعة الأنشطة، والمقدمة بأساليب علمية في وثائق رسمية أو غير رسمية، لخدمة غرض محدد.

و يظهر إذن أن عملية التعداد هي أداة لتوفير المعطيات الإحصائية وبتعبير آخر هي أداة لتوفير الإحصائيات.

إن جميع المعطيات العددية لمختلف الظواهر، حسب هذا المفهوم يعود الى أقدم العصور، حيث كانت تنظيمات القبائل والعشائر، تلجأ لأسلوب العد سواء لتحديد عدد السكان الخاضعين للضريبة أو الخدمة العسكرية، أو تحديد الأملاك كالماشية أو الإبل.. الخ. غير أن عملية الإحصاء كانت عموما من مهام الحكومة، وقد أنشئت لأجلها أجهزة خاصة تطورت مع الزمن، ومع تطور العلوم، من إستخدام الوسائل البدائية الى استخدام التقنيات العالية، كأجهزة الإعلام الآلي، لتخزين ومعالجة المعلومات المحصل عليها، وأصبح الإقتصادي أو الباحث الذي يريد إستخدام الإحصائيات في بحوثه من السهل عليه الحصول عليها من منشورات الهيئات الرسمية دون الحاجة الى إجراء جمع المعطيات الإحصائية بنفسه.

في الجزائر نجد أن الهيئة الوطنية التي تقوم بجمع المعطيات الإحصائية ومعالجتها من مختلف الأنشطة الإقتصادية والاجتماعية عبر الوطن، هي الديوان الوطني للإحصائيات، وهو الهيئة الرسمية الأساسية المنتشرة عبر ربوع الوطن من

خلال ملحقاتها الجهوية في كل من الجزائر، وهران، بشار،
قسنطينة، عنابة وورقلة.

نبذة حول الديوان الوطني للإحصائيات كما يقدمها عن نفسه

(أنظر: www.ons.dz/st1.htm)

الديوان الوطني للإحصائيات، مؤسسة مركزية للإحصائيات بالجزائر و هي ذات طابع إداري مكلفة بجمع و معالجة و نشر المعلومات الإحصائية الاجتماعية و الاقتصادية مثل إحصاء السكان و السكن، مسح حول المؤسسات الصناعية .. الخ، الديوان الوطني للإحصائيات تحت وصاية الوزير المنتدب لدى رئيس الحكومة المكلف بالتخطيط.

أنشئ الديوان سنة 1964 تحت اسم المحافظة الوطنية لإحصاء السكان، و هذا لغرض القيام بإنجاز الإحصاء الأول للسكان في الجزائر المستقلة سنة 1966، و في سنة 1971 تم تغيير تسميته ليصبح المحافظة الوطنية للإحصائيات و المسوحات الإحصائية، و قد تم إنجاز أعمال معتبرة خلال هذه الفترة مثل الإحصاء الثاني للسكان و السكن سنة 1977 و المسح الديموغرافي خلال 1972-1973 و المسح الخرائطي خلال 1972-1975، و الذي يعتبر قاعدة لإنجاز الإحصاء و مسح إستهلاك الأسر في 1979-1980.

و تمت إعادة هيكلة الجهاز و إنشاء الديوان الوطني للإحصائيات الحالي بمقتضى المرسوم التشريعي رقم 82-484 المؤرخ في 18 ديسمبر 1982 و المعدل بموجب المرسوم رقم 85-311 بتاريخ 17 ديسمبر 1985، و قد أعيد تنظيمه بموجب المرسوم 95-159 بتاريخ 3 جويلية 1995.

من مهام الديوان الوطني للإحصائيات، أنه يسهر على إعداد و توفير و نشر المعلومات الموثوق بها و المنتظمة و المتماشية و حاجيات الأعوان الاقتصاديين و الاجتماعيين، و يعمل على ضمان التوفير المنتظم للمعلومات و التحاليل الإحصائية و الدراسات الاقتصادية الضرورية لإعداد و متابعة السياسات الاقتصادية و الاجتماعية التي تعدها السلطات العمومية، فهو يعد و ينشر الأرقام الاستدلالية و مؤشرات الاقتصاد الوطني و حسابات الأمة، و يسير التسجيلات الإحصائية و المسوحات و الأعمال الإحصائية و يعد و يحين فهرس الأعوان الاقتصاديين و الاجتماعيين و هذا بتخصيص رقم تعريف إحصائي لكل عون.

ومن أضخم انجازات الديوان هو تعداد السكان، وكان أهم عمل قام به بهذا الخصوص هو التعداد العام لسنة 1977، وكذلك الذي أجراه سنة 1987، و سنة 1998، كما يقوم كذلك بإجراء وجمع احصائيات حول معظم النشاطات الاقتصادية والاجتماعية، كالتشغيل، الصحة، التعليم، الفلاحة، الصناعة، النقل، البريد والمواصلات، التجارة الخارجية، الأسعار، المالية، الحسابات الوطنية... الخ. كما يقوم الديوان بمعالجة هذه البيانات ونشرها في مطبوعاته الخاصة، أهمها المجموعة الإحصائية السنوية، وهي عبارة عن مطبوع من الحجم الكبير يتضمن معطيات إحصائية عن كل الأنشطة الاقتصادية والاجتماعية للوطن، و من ذلك السكان والسكن، نشاطات الصحة و التربية و التعليم و التكوين، الأنشطة الاقتصادية كالصناعات و التجارة الداخلية و الخارجية، الأسعار والمداخيل، و الشؤون المالية و الحسابات الوطنية، كما يصدر هذه الإحصائيات ملخصة أيضا من خلال مجلة الجزائر بالأرقام، و يقوم ببعض التحليلات من خلال مجلة إحصائيات، التي ينشر فيها أيضا بعض الدراسات التي يقوم بها، كما يصدر أيضا بعض الكتيبات والدوريات الأخرى التي تهتم إما بنشر ملخصات إحصائية أو دراسات حول ظاهرة ما، أو مجموعة من الظواهر.

ان جمع المعطيات الإحصائية ونشرها وحده لا يكفي، إذ لابد من إستخدام هذه المعطيات في مختلف الميادين التحليلية، لغرض حصر الإمكانيات، والاستخدام الأمثل لها، في مجالات التخطيط والتنبؤ واتخاذ القرارات، وذلك بأساليب وفتيات معينة يتناولها علم الإحصاء.

3- **علم الإحصاء:** يختلف مصطلح الإحصاء عن الإحصائيات بالمفهوم الذي أعطيناه آنفاً، و عموماً يقصد بلفظ الإحصاء في المجال الأكاديمي علم الإحصاء و هو علم متطور باستمرار كغيره من العلوم، فما هو علم الإحصاء إذن؟.

تعريفه 1-2: الإحصاء هو علم يدرس مختلف طرق و وسائل جمع البيانات الكمية عن مختلف الظواهر الاقتصادية والاجتماعية... وترتيبها وتبويبها وتحليلها وتفسيرها، وتقديمها بأشكال وصور ملائمة بهدف تسهيل اتخاذ القرار على أساس سليم.

أي أنه يهدف الى دراسة كيفية جمع المعلومات الكمية عن الظواهر الاقتصادية والاجتماعية و الطبيعية...، وكيفية تحويلها الى جداول وأشكال معينة، لتسهيل فهمها واستخدامها، كما يهدف أيضاً الى تقديم الأساليب المختلفة التي تستخدم لتحليل هذه المعلومات سواء كانت هذه الأساليب وصفية أو رياضية، وبناء على هذا فإن الموضوعات التي يتناولها علم الإحصاء هي :

- طرق ووسائل جمع المعلومات الكمية عن الظواهر.
- تصنيف تلك المعلومات و ترتيبها.
- تقديم تلك المعلومات بأشكال وصور ملائمة.
- تحليل المعلومات المحصل عليها، واستقراء النتائج، واستخدامها في مجالات التنبؤ و التحضير لاتخاذ القرارات.
- فعلم الإحصاء اذن بهذا المفهوم علم واسع، بطرقه المختلفة و قوانينه المتعددة وأساسه الثابتة و نظرياته العلمية و تطبيقاته الواسعة الانتشار، وله علاقات متشعبة و متبادلة مع مجمل العلوم الأخرى، حيث يؤثر فيها ويتأثر بها.

إن إضفاء الصفة العلمية على الإحصاء راجع لكونه يتميز بكل مواصفات العلوم الأخرى، إذ يستخدم طرق البحث

العلمي المعروفة كالطريقة الإستقرائية أو الطريقة الإستنباطية، كما أن له موضوعا وهدفا محددًا، يتم الوصول اليه بإتباع المنهجية العلمية، إضافة الى أن له قوانين و نظريات ثابتة وصالحة في كل زمان ومكان، كما أنه يستخدم من طرف الكثير من العلوم، بل و هناك من العلوم ما يعتبر الإحصاء مقومها الرئيسي، و بدوره تعتبر الرياضيات أدواته الأساسية.

تقليديا تدخل المواضيع التي تتعلق بدراسة طرق ووسائل جمع البيانات الكمية وتصنيفها و تبويبها ضمن جداول إحصائية وكذا تقديم وسائل تحليلها الوصفية، ضمن ما يسمى بالإحصاء الوصفي، الذي هو موضوع هذا الكتاب، بينما تدخل عمليات التحليل المعمق للبيانات الإحصائية واستقراء النتائج و إختبار الفرضيات و الاحتمالات الى غير ذلك من المواضيع، ضمن ما يسمى بالإحصاء الرياضي أو الطرق الإحصائية، وعمليا يعتبر الإحصاء الوصفي كمدخل أساسي لعلم الإحصاء الواسع، لكونه يقدم المبادئ الأولى التي يجب على الإحصائي الإلمام بها، كي يتمكن منولوج في أعماق الإحصاء و يتمكن بالتالي من إجراء دراسات إحصائية علمية معمقة.

ومن الأدوات الهامة التي تدخل ضمن وسائل وصف الظواهر والتي تدخل ضمن الإحصاء الوصفي و التي سنتناولها في هذا الكتاب ما يلي :

- العرض الجدولي و البياني للمعلومات الإحصائية.

- مقاييس التزعة المركزية.

- مقاييس التشتت.

- أشكال التوزيعات.

- الإنحدار والارتباط.

- السلاسل الزمنية.

- الأرقام القياسية و معدلات النمو.

ثانياً: مجالات استخدام الإحصاء: يستخدم علم الإحصاء في معظم مجالات الحياة اليومية، وفي مجالات البحوث العلمية المختلفة، فعلم الاقتصاد، يستخدمه في تحليل الظواهر وإبراز العلاقات بينها وفي مجال التخطيط وغير ذلك، و العلوم الطبية والبيولوجيا بصفة عامة، تستخدمه في حصر نتائج التجارب ودراساتها... وكذلك بالنسبة للعلوم الاجتماعية و الادارية، فهي تعتمد اعتماداً واسعاً على الإحصاء. غير انه ينبغي الإشارة الى أن علم الإحصاء يعمل دائماً في الاتجاه الذي يحدده له العلم الذي يستخدمه.

ثالثاً: أنواع البحوث الإحصائية : وفقاً للتعريف السابق للإحصاء، فإنه يمكن تقسيم البحوث الإحصائية الى ثلاثة أنواع هي :

1- البحوث الإحصائية الوصفية : وهي البحوث التي تجمع فيها البيانات عن الظواهر، بهدف توفيرها وعرضها لتستخدم لأغراض معينة من طرف باحثين آخرين، فهدف جمع البيانات في هذا النوع من البحوث لا يكون محددًا سلفاً، ومن أمثلة ذلك، البحوث التي يقوم بها الديوان الوطني للإحصائيات، كالتعداد العام للسكان وإحصائيات التجارة الخارجية، والأنشطة الصناعية وغيرها.

2- البحوث الإحصائية التحليلية : وهي البحوث التي تجمع فيها البيانات عن الظواهر لأجل خدمة هدف محدد من طرف الباحث، سواء كان الباحث شخصاً طبيعياً أو اعتبارياً، وسواء قام هو نفسه بجمع البيانات الإحصائية من مصادرها المباشرة أو اعتمد على إحصائيات معدة من طرف باحث آخر أو هيئة إحصائية، و بمعنى آخر هي البحوث التي تهدف الى تحليل ظاهرة ما للوصول الى هدف محدد سلفاً، ومن أمثلة ذلك، البحوث

التي تجرى لغرض معرفة مستوى التغذية في إحدى الولايات، أو البحوث التي تهدف الى معرفة مستوى البطالة في مجتمع ما... الخ.

3- البحوث الإحصائية التجريبية : وهي البحوث التي تجرى لغرض محدد سلفا، فهي بذلك بحوث تحليلية من جانب، غير أن ظروف إنجازها تخضع للباحث نفسه، فهو يتحكم في توفير الظروف المساعدة لإنجاز البحث، وهذا هو الفرق بينها وبين البحوث التحليلية، و تجرى هذه البحوث في غالب الأحيان في الدراسات الفلاحية و البيولوجية عامة.

رابعاً: منهجية البحث الإحصائي : لأجل إجراء أي عمل إحصائي لابد من اتباع منهجية منطقية تؤدي بالباحث في النهاية الى إجراء العمل الإحصائي في أقصر وقت وبأقل تكلفة وجهد ممكنين. ويمكن تلخيص محاور هذه المنهجية في النقاط الرئيسية المرتبة في شكل مراحل كمايلي:

المرحلة 1 : تحديد الظاهرة المدروسة.

المرحلة 2: جمع البيانات الإحصائية.

المرحلة 3 : تصنيف وتبويب البيانات.

المرحلة 4: تحليل البيانات إحصائياً، و استقراء النتائج.

1 - المرحلة الأولى: التحديد الدقيق للظاهرة المدروسة : أول مرحلة في البحث الإحصائي، هي التحديد العام للظاهرة المدروسة، إذ على الباحث أن يحدد بكل دقة الهدف من الدراسة الإحصائية، ثم المجتمع الإحصائي و مكانه و الوقت المناسب لجمع البيانات حوله، والصفات المطلوب معرفتها ووحدات القياس المستخدمة.

1 - تحديد الهدف، مكان و زمان دراسة الظاهرة : إن تحديد الهدف من الدراسة الإحصائية و المكان و الزمان

المناسب لجمع البيانات هو أول ما يجب على الباحث أن يحدده بكل دقة، وذلك من خلال الإجابة على الأسئلة التالية:

- لماذا يتم إجراء هذه الدراسة، أي ماهو الغرض أو الهدف من جمع البيانات الإحصائية حول الظاهرة ؟
- أين توجد الظاهرة المدروسة ؟ أي التحديد المكاني لها.

- ماهي الفترة الزمنية المناسبة لأجراء عملية جمع البيانات؟ صباحا أم مساء، صيفا أم شتاء... ؟.

فإذا ما كانت الظاهرة المدروسة مثلا، هي مدى كفاية المنحة الوطنية لاحتياجات الطالب الشهرية، فإن الهدف المراد الوصول اليه هو ناتج الاجابة على السؤال التالي : هل المنحة المقدمة للطالب كافية لمصاريفه المتعلقة بالدراسة أم لا؟ وما هي هذه المصاريف ؟ و ماهي قيمها؟ وبعد تحديد الهدف علينا أن نحدد مكان وجود الظاهرة المدروسة، و مادام الامر يتعلق بدراسة مدى كفاية المنحة الدراسية للطالب، لذلك فان المكان الذي تجرى فيه عملية جمع البيانات هو الجامعة، وعلينا ان نحدد بعد ذلك أي جامعة تجرى فيها العملية، هل في وسط البلاد أم في غربها أم تجرى في جميع جامعات البلاد ؟ ثم أخيرا علينا أن نحدد العناصر التي ينبغي أن نجري عليها العملية وأن نحدد وحدات قياس الكميات، هل تقاس بالوحدات النقدية مثلا، أم بوحدات الأوزان أم بغير ذلك ؟. إن الاجابة على جميع هذه الاسئلة، تفيد الباحث في مجالات عديدة منها:

- عدم بذل الجهود في سبيل الحصول على معلومات غير مفيدة للدراسة وبالتالي عدم إضاعة الوقت والجهد و المال.
- عدم اهمال الحصول على بعض المعلومات التي قد تكون هامة وضرورية للدراسة.

- التمكن من حصر الوسائل المادية والبشرية الضرورية لأجل إجراء الدراسة.

بج - المجتمع الإحصائي، الوحدة الإحصائية، الصفة:
إن تحديد المجتمع الإحصائي أمر في غاية الأهمية، ويعرف المجتمع الإحصائي بأنه مجموع الوحدات الإحصائية التي تقع عليها الدراسة الإحصائية، فإذا كان الهدف هو إجراء عملية إحصائية حول النفقات اليومية للطالب في جامعة ما مثلاً، فإن المجتمع الإحصائي هو مجموع طلبة هذه الجامعة، والوحدة الإحصائية هي الطالب الفرد، فالوحدة الإحصائية هي العنصر الأول محل الدراسة الإحصائية، أو هي القيمة المادية أو المعنوية التي تقع عليها الدراسة الإحصائية.

و الوحدة الإحصائية لظاهرة ما تكون من نوع معين ذي صفات مشتركة نستطيع عن طريقها تصنيفها في مجموعة إحصائية واحدة هي المجتمع الإحصائي، وقد تحتوي الوحدة الإحصائية عنصراً واحداً أو مجموعة من العناصر المشتركة بينها يتم الإستقصاء عنها، ويسمى هذا العنصر أو العناصر بالصفة أو الصفات، فعند الدراسة الإحصائية لمجتمع طلابي مثلاً، فإنه يمكن التكلم عن مجموعة من الصفات المشتركة، ومنها الفرع الذي ينتمي إليه الطالب، التخصص، الجنس، ... الخ، و تسمى مثل هذه الصفات بالصفات النوعية، و إذا ما كانت الدراسة الإحصائية تتعلق بمردودية مجموعة من المؤسسات مثلاً، فإن الصفات المشتركة التي يمكن التكلم عنها هي حجم الإنتاج فيزيائياً أو نقدياً، تكاليف الإنتاج، الأرباح... الخ، و مثل هذه الصفات تسمى بالصفات الكمية.

وعموماً نقول بأن للوحدة الإحصائية صفة نوعية، إذا لم تكن هذه الصفة قابلة للقياس كالجنس أو الرتبة أو التخصص... الخ،

و نقول بأن للوحدة الإحصائية صفة كمية إذا كانت هذه الصفة قابلة للقياس عن طريق وحدة من وحدات القياس. و قد تكون وحدات القياس أشياء معدودة كعدد الطلبة أو التجار مثلاً، أو وحدات قياس طبيعية أو فيزيائية سواء كانت بسيطة كالكيلوغرام والمتر وأجزائهما وأضعافهما أو مركبة كالتر\الثانية أو اللتر\الثانية، وقد تكون وحدات نقدية كالدينار أو الدولار. ويمكن لعناصر الوحدة الإحصائية أن تجمع بين الصفتين الكمية والنوعية في آن واحد.

2- المرحلة الثانية: جمع البيانات الإحصائية : إن جمع البيانات الإحصائية من أساسيات العمل الإحصائي، ولهذه المرحلة أهمية خاصة، في أي بحث إحصائي، إذ أن توفر البيانات الإحصائية الدقيقة والسليمة عن الظاهرة المدروسة، يعطي نتائج سليمة، ويساعد على اتخاذ قرار سليم بناء على تلك النتائج، وعلى الباحث أن يحدد مصدر جمع البيانات المرغوب فيها، وأساليب وطرق ذلك، قبل البدء في العملية.

أ - مصادر جمع البيانات : مصادر جمع البيانات هي منابع التي يأخذ منها الإحصائي البيانات موضع الدراسة، وقد تكون هذه المصادر مباشرة وقد تكون غير مباشرة.

• **مصادر مباشرة:** في هذه الحالة يجمع الإحصائي معلوماته إما بالاتصال و الاحتكاك المباشر بوحدات المجتمع الإحصائي أو بالاعتماد على الوثائق التي تكون فيها المعلومات لازالت خاماً.

• **مصادر غير مباشرة :** في هذه الحالة يتم الحصول على البيانات من مصادر غير الأولية، حيث تكون مبوبة ومصنفة من قبل باحثين سابقين أو هيئات رسمية، أو غير رسمية،

وتم نشرها في نشرات خاصة أو دوريات، أو تكون محفوظة في الأرشيف التقليدي أو الآلي.

ب- أساليب جمع البيانات من المصادر المباشرة :
يمكن جمع البيانات الإحصائية بأحد الأسلوبين التاليين:

*** أسلوب الحصر الشامل :** في هذه الحالة، تتم دراسة كل وحدات المجتمع الإحصائي أي أخذ المعلومات المراد الحصول عليها مباشرة من الوحدة الإحصائية، ومن مميزات هذا الأسلوب أنه يوفر حظوظ الحصول على معلومات دقيقة إذا ما توفرت شروط البحث الإحصائي، الشيء الذي يجعل نتائج الدراسة الإحصائية غير مشكوك فيها، غير أن أهم عيوبه إرتفاع التكاليف لكون عملية الحصر الشامل تتطلب وسائل مادية وبشرية ضخمة، كما أن هذا الأسلوب لا يتماشى مع الدراسات التي يلعب الوقت فيها دورا حاسما، كما أنه كلما كان حجم المجتمع كبيرا كلما كان احتمال الخطأ كبيرا، ونظرا لكون هذا العمل شاملا ويتطلب وسائل مادية و بشرية ضخمة، لذلك فهو في غالب الأحيان من مهمة الحكومات خاصة إذا ما كان يتعلق بدراسات سوسيولوجية.

*** أسلوب العينات :** العينة تعريفا هي جزء من المجتمع المراد دراسته، وتحدد بعدة طرق، منها طريقة العينة العشوائية البسيطة، وطريقة العينة الطبقية، والعنقودية و المنتظمة. ومن الأسباب الرئيسية لإنتهاج هذا الأسلوب مايلي :

- أن العينة يمكن أن تمثل كل المجتمع الإحصائي من حيث المميزات والخصائص.

- الإقتصاد في الوقت والجهد و المال.

- إمكانية فحص جميع وحدات المجتمع الإحصائي، فمثلا عندما يريد الباحث معرفة المسواد أو المعنادن التي تتكون منها تربة

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018
منطقة ما، فإنه يستحيل عليه فحص كل التربة، و يكفي لذلك
أن يأخذ كميات من التربة من مناطق مختلفة ويفحصها
ويستخرج النتيجة.

- كما أنه في بعض الحالات تؤدي دراسة الوحدة
الإحصائية إلى اتلافها، فعند دراسة مكونات حبات البيض مثلاً
في حضيرة ما، فإن دراسة كل المتوج تؤدي إلى اتلافه كلية،
لذلك يلجأ إلى دراسة عينة منه فقط، وفق معايير محددة،
وتعميم النتائج على جميع وحدات المجتمع، و في التحاليل الطبية
يلجأ المختبر إلى أخذ عينات فقط من دم الإنسان لتحليلها لأن
سحب كامل كميات الدم منه لتحليلها تؤدي إلى هلاكه.
- عدم إمكانية إجراء أسلوب الحصر الشامل، فيكون
الباحث ملزماً باستخدام أسلوب العينات كما في حالة دراسة
الأسماك أو الطيور أو الحيوانات الغريبة.

- وجود قيود الزمن و التكاليف المخصصة لإنجاز عملية الإحصاء،
إذ قد لا تسمح هذه القيود بإجراء العملية بأسلوب المسح الشامل.
و تختار العينة بمجموعة من الطرق منها ما يلي:

**** العينات العشوائية: و هناك عدة طرق**

لإختيارها منها ما يلي:

- **طريقة العينة العشوائية البسيطة:** في هذه الطريقة
تؤخذ العينة بشكل يعطي لأي عنصر من عناصر المجتمع نفس
الفرصة لأن يكون ضمن العينة، وتؤخذ العينة بإحدى
الطريقتين، إما أن يتم خلط وحدات المجتمع الإحصائي خلطاً
جيداً، ويتم أخذ وحدات العينة بصفة عشوائية، بحيث يكون
لكل وحدة نفس احتمال الظهور، وإما باستخدام جداول
الأرقام العشوائية و يتم ذلك كما يلي:

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

نعطي كل عناصر المجتمع الإحصائي أرقاماً متسلسلة من 0 إلى N-1، حيث N عدد وحدات المجتمع، وبحيث يكون لكل رقم من عناصر المجتمع أعداداً تساوي عدد أرقام حجم المجتمع، فإذا كان حجم المجتمع 53 مثلاً، فإن كل وحدة في المجتمع الإحصائي تعطى رقماً مكوناً من عددين بدءاً من : 00 01 02 10 11 52، وإذا كان حجم المجتمع 625 مثلاً، فإن كل وحدة من وحدات المجتمع تأخذ رقماً مكوناً من 3 أعداد بدءاً من : 000 001 002 ... 010 011 ... 099 100 101 623 624. وهكذا...

نأخذ صفحة من صفحات الجداول العشوائية (أنظر الصفحة الموالية)، ونختار عموداً يكون عدد أرقامه مساوياً لعدد أرقام حجم المجتمع، ونأخذ كل الأرقام المحصورة ضمن المجال ونلغي البقية.

مثال 1-1: نريد إختيار عينة تتكون من 10 طلبة من مجتمع طلابي مكون من 400 طالب بطريقة العينة العشوائية البسيطة، وباستخدام جداول الأرقام العشوائية. بين كيف يتم ذلك؟.

نعطي كل طالب رقماً من الأرقام التالية : 000 001 002 003 ... 009 010 011 012 013 ... 020 021 022 023 ... 099 100 101 102 103 ... 200 201 203 .. 300 301 302 303 ... 397 398 399.

نأخذ صفحة من صفحات جداول الأرقام العشوائية، و نعين عمود يتكون من 3 أرقام ونأخذ على التوالي 10 أرقام من الأرقام التي نصادفها ضمنه والتي تكون محصورة بين 000 و 399، ويكون الطالب الذي يحمل رقماً من الأرقام العشرة المأخوذة من ضمن عناصر العينة المختارة.

صفحة من جداول الأرقام العشوائية

31 75 15 72 60 68 98 00 53 39 15 47 04 83 55 88 65 12 25 96 03 15 21 91 21
88 49 29 93 82 14 45 40 45 04 20 09 49 89 77 74 84 39 34 13 22 10 97 85 08
30 93 44 77 44 07 48 18 38 28 73 78 80 65 33 28 59 72 04 05 94 20 52 03 80
22 88 84 88 93 27 49 99 87 48 60 53 04 51 28 74 02 28 46 17 82 03 71 02 68
78 21 21 69 93 35 90 29 13 86 44 37 21 54 86 65 74 11 40 14 87 48 13 72 20

41 84 98 45 47 46 85 05 23 26 34 67 75 83 00 74 91 06 43 45 19 32 58 15 49
46 35 23 30 49 69 24 89 34 60 45 30 50 75 21 61 31 83 18 55 14 41 37 09 51
11 08 79 62 94 14 01 33 17 92 59 74 76 72 77 76 50 33 45 13 39 66 37 75 44
52 70 10 83 37 56 30 38 73 15 16 52 06 96 76 11 65 49 98 93 02 18 16 81 61
57 27 53 68 98 81 30 44 85 85 68 65 22 73 76 92 85 25 58 66 88 44 80 35 84

20 85 77 31 56 70 28 42 43 26 79 37 59 52 20 01 15 96 32 67 10 52 24 83 91
15 63 38 49 24 90 41 59 36 14 33 52 12 66 65 55 82 34 76 41 86 22 53 17 04
92 69 44 82 97 39 90 40 21 15 59 58 94 90 67 66 82 14 15 75 49 76 70 40 37
77 61 31 90 19 88 15 20 00 80 20 55 49 14 09 96 27 74 82 57 50 81 69 76 16
38 68 83 24 86 45 13 46 35 45 59 40 47 20 59 43 94 75 16 80 43 85 25 96 93

25 16 30 18 89 70 01 41 50 21 41 29 06 73 12 71 85 71 59 57 68 97 11 14 03
65 25 10 76 29 37 23 93 32 95 05 87 00 11 19 92 78 42 63 40 18 47 76 56 22
36 81 54 36 25 18 63 73 75 09 82 44 49 90 05 04 92 17 37 01 14 70 79 39 97
64 39 71 16 92 05 32 78 21 62 20 24 78 17 59 45 19 72 53 32 83 74 52 25 67
04 51 52 56 24 95 09 66 79 46 48 46 08 55 58 15 19 11 87 82 16 93 18 33 61

15 88 09 22 61 17 29 28 81 90 61 78 14 88 98 92 52 52 12 83 88 08 16 00 98
71 92 60 08 19 59 14 40 02 24 30 57 09 01 94 18 32 90 69 99 26 85 71 92 38
64 42 52 81 08 16 55 41 60 16 00 04 28 32 29 10 33 33 61 68 65 61 79 48 34
79 78 22 39 24 49 44 03 04 32 81 07 73 15 43 95 21 66 48 65 13 65 85 10 81
36 33 77 45 38 44 55 36 46 72 90 96 04 18 49 93 86 54 46 08 93 17 63 48 51

(إذا لم يكفي عمود واحد لإختيار العينة تنتقل الى عمود آخر).
تجدر الإشارة الى أن العينة العشوائية البسيطة يمكن أن تسحب بأحد الأسلوبين، إما السحب بدون إعادة، أي رفض الرقم الذي سبق وأن أخذ، أو السحب مع الإعادة، أي قبول مرة أخرى الرقم الذي سبق وأن أخذ.

- **طريقة العينة الطبقية:** تستخدم هذه الطريقة في الحالات التي تكون فيها نتيجة البحث تعتمد على العمر أو الجنس أو المكان أو الدخل... الخ، ويتم في هذه الحالة تقسيم المجتمع الإحصائي الى مجموعات جزئية تعتمد على هذه الصفات وتسمى بالطبقات، ثم باستخدام طريقة العينة العشوائية البسيطة يتم إختيار عينة جزئية من كل طبقة يتناسب حجمها مع حجم الطبقة، وتشكل مجموعة العينات الجزئية المختارة، ما يسمى بالعينة الطبقية.

تستخدم العينة الطبقية لعدة أسباب منها ما يلي :

- قد تكون هذه الطريقة مناسبة من الناحية الإدارية.
- قد يكون هناك تمايزا واضحا بين طبقات المجتمع الإحصائي.

- قد تكون خاصية من خواص المجتمع تختلف اختلافا كبيرا من منطقة لأخرى أو من فئة لأخرى.

- **طريقة العينة العنقودية:** في بعض الأحيان يرى الإحصائي بأن الطريقتين السابقتين غير مناسبتين لإختيار العينة، فيلجأ الى طريقة العينة العنقودية، حيث يقوم بتقسيم المجتمع الإحصائي الى مجموعات جزئية واضحة، ثم تختار من كل مجموعة جزئية مجموعة جزئية أخرى أقل منها تسمى عنقودا، ثم تختار من هذه العناقيد بطريقة العينة العشوائية البسيطة عينات

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018
جزئية، بحيث تشكل مجموعة العينات الجزئية هذه ما يسمى
بالعينة العنقودية.

- **طريقة العينة المنتظمة:** في هذه الطريقة تكون
قاعدة إختيار عناصر العينة وفق نظام معين، فإذا ما أردنا
التعرف على مدى كفاية المنحة الدراسية للطلبة بإستخدام هذه
الطريقة مثلاً، نقوم بأخذ مكان عند المدخل الرئيسي للجامعة،
ونستجوب كل عاشر أو كل خامس طالب يدخل الى الحرم
الجامعي، أو أن نعطي الطلبة أرقاماً من 0 الى N-1، ونختار
رقماً عشوائياً ضمن هذا المجال، كأن نختار الرقم 5 مثلاً بصفة
عشوائية، ونضيف في كل مرة رقماً ثابتاً بصفة تسلسلية وتكون
الأرقام المحصل عليها هي أرقام الطلبة الذين يشكلون العينة،
فإذا أردنا أن يكون حجم العينة 12 طالب واخترنا و كان الرقم
الأول المختار عشوائياً هو 8 مثلاً، و الرقم المضاف بانتظام هو
10 مثلاً، فإن أرقام الطلبة الذين يشكلون العينة هم :

8 18 28 38 48 58 68 78 88 98 108 118

و يلاحظ أن هذه الأرقام مأخوذة بانتظام و هذا ما جعلنا نسمي
هذه الطريقة بطريقة العينة المنتظمة.

•• **العينات غير العشوائية:** و هي التي لا تخضع
لقانون العشوائية، إنما يتم إختيارها بانحياز، لإعتبرات تعود
للباحث الإحصائي و منا ما يلي:

- **العينة السهلة المنال:** في هذا النوع من العينات
يكون المعيار الوحيد المتخذ هو سهولة حصول الباحث على
مفردات العينة، تستخدم في الغالب في البحوث الإستطلاعية
التي لا تتصف بالدقة الكاملة، لكنها تتميز بسرعة الحصول
عليها و قلة تكاليف الحصول عليها، لكونها مأخوذة بتحيز فإنه

تحتوي على الكثير من الأخطاء، و على الباحث أن يراعي ذلك عند محاولة تفسير و تعميم نتائجها.

- **العينة الحصية:** و يتم اختيارها أيضا بتحيز قصد إظهار الخصائص ذات الأهمية للباحث الإحصائي، غير أنه ينبغي توافر الشروط التالية عند اللجوء لإستخدامها و ن ذلك ما يلي:

- أن تكون الخصائص الخاصة بالمجتمع و التي تؤثر على موضوع البحث متوافرة في العينة.
- أن تكون هذه الخصائص من الممكن إستخدامها في تقسيم المجتمع الى مجموعات متجانسة كالدخل أو السن أو الوظيفة... الخ.
- أن تكون هذه الخصائص مؤثرة تأثيرا ملحوظا على موضوع البحث.
- أن يكون عدد هذه الخصائص محدودا حتى لا توجد مجموعات كثيرة يصعب التعامل معها.

- **العينات العمدية:** وهي التي يتم اختيارها و هناك بعض الأهداف المحددة في ذهن الباحث الذي يقوم باختيار العينة، و هذه العينة لا تمثل المجتمع الإحصائي، لذا ينبغي عليه توخي الحذر في تفسيره لنتائج دراسة العينة.

- **العينات العكسية:** وهي التي يقوم الباحث باختيارها بصورة تمثل المجتمع مع إستخدامه لبعض المعايير الحكمية القائمة على خبرته الشخصية، و بالتالي فإن جودة هذه العينة تتوقف بدرجة كبيرة على خبرة الباحث الذي يقوم بعملية الاختيار.

و لاشك أن طرق العينات العشوائية تكون الأكثر مصداقية في مختلف الدراسات الإحصائية لكونها خالية من عنصر التحيز.

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

ج - طرق جمع البيانات الإحصائية: تجمع البيانات الإحصائية حسب طبيعة المجتمع الإحصائي، إما بالاتصال المباشر بوحدة المجتمع الإحصائي أو العينة الإحصائية، أو بالاتصال غير المباشر بهما، أو عن طريق الملاحظة أو التسجيل الحيوي.

• **طريقة الإتصال المباشر:** في هذه الحالة يتطلب الأمر الإتصال بجميع وحدات المجتمع الإحصائي أو العينة، وذلك إما عن طريق الباحث نفسه وذلك عندما يكون حجم المجتمع الإحصائي صغيراً، بحيث يمكن للباحث أن يتحكم فيه من حيث جمع البيانات، ومن حيث التكاليف والوقت، وفي هذه الحالة تكون البيانات المحصل عليها في غاية الدقة نتيجة لإهتمام وحرص الباحث عليها لكونها تهمه مباشرة، وبالتالي تكون نتائج الدراسة التي يتوصل إليها صحيحة، ومن عيوب هذه الطريقة أنها تستغرق وقتاً طويلاً. وقد يتم استخدام أعوان إحصائيين، ويتم ذلك عندما يكون حجم المجتمع أو العينة كبيراً، بحيث يستحيل على الباحث أن يقوم وحده بذلك، غير أنه ينبغي إعداد الأعوان إعداداً جيداً، كما يجب أن يتصفوا باللياقة والدبلوماسية وكرم الأخلاق، وأن يتحلوا بالصبر والأدب والأمانة في تدوين المعلومات، إضافة إلى ضرورة إلمامهم بكل المعاني والمصطلحات والتعاريف التي تتضمنها الإستمارة الإحصائية، كما يجب أن يكونوا قادرين على التفسير والإقناع بما يمكن أن يوجه إليهم من أسئلة واستفسارات، ومن مزايا هذه الطريقة أنها تسمح بتفسير الأسئلة وتوضيح كل ما يمكن أن يكون غامضاً، كما تسمح بجمع المعلومات في أقصر وقت ممكن، ولها عيوباً كثيرة أهمها إرتفاع التكاليف وإحتمال الوقوع في أخطاء كثيرة، إضافة إلى

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

إحتمال التأثير السلبي للأعوان إذا لم يتحلوا بالصفات المشار إليها، مما يدفع المبحوثين للإدلاء بمعلومات غير صحيحة، كما أن هذه الطريقة غير صالحة إذا ما كانت الأسئلة تتعلق بأمور شخصية بحتة، كالأمور العائلية أو السياسية أو المذهبية.

• **طريقة الإتصال غير المباشر:** في هذه الطريقة يتم جمع البيانات الإحصائية دون أن يتصل الباحث أو الأعوان مباشرة بوححدات المجتمع الإحصائي أو العينة، وذلك إما عن طريق المراسلة أو الهاتف أو الأنترنت أو وسائل الإعلام، و تستخدم في المجتمعات التي تتميز بالوعي الإحصائي.

•• **طريقة المراسلة:** في هذه الطريقة يقوم الباحث بارسال إستمارة إلى عناصر المجتمع الإحصائي أو العينة، التي يجري العمل الإحصائي عليها، حيث يُطلب ملأها و إعادتها إلى الباحث، عن طريق المراسلة أيضا، ولهذه الطريقة عدة مزايا أهمها قلة التكاليف، و إقتصاد الوقت، ومن عيوبها عدم الدقة في الإجابات ولامبالاة بعض المبحوثين بها.

•• **المخاطبة الهاتفية:** تصلح هذه الطريقة في الدراسات المحدودة والتي يضمن فيها الباحث وجود أجهزة هاتفية لدى المبحوثين، ومن مزاياها أنها توفر الوقت بحيث تسمح للباحث الحصول على المعلومات التي يرغب في الحصول عليها بسرعة فائقة بعيدا عن سوء فهم الأسئلة، كما أن المعلومات المحصل عليها تكون دقيقة إضافة إلى انخفاض تكاليف العملية.

•• **من طرق وسائل الإعلام:** ويتم ذلك بطرح أسئلة محددة، عبر وسائل الاعلام كالصحف، الإذاعات، والتلفزيون، الأنترنت، بحيث تتم الإجابة عنها إما عن طريق الهاتف مباشرة، أو عن طريق المراسلة، ومن مزاياها أنها قليلة التكاليف

و يمكن أن تصل جميع عناصر المجتمع الإحصائي، وأهم عيوبها إمكانية لامبالاة المبحوثين بها.

• طريقة الملاحظة والرقابة: بعض الدراسات أو

البحوث تتطلب إتباع رصد الظاهرة عن طريق ملاحظتها وتسجيل قياساتها وتحرك متغيراتها، ويتم ذلك خاصة في البحوث البيولوجية و الدراسات العملية، و في بعض الدراسات الاجتماعية، حيث يقوم الباحث برصد الحركة والوقت والتعرف على محددات الظاهرة، كأن يقوم برصد حركة المرور عندما تكون الدراسة تتعلق بذلك، عند فترات زمنية محددة، ليتم إستنتاج الأوقات التي يكثر فيها الإزدحام والأوقات التي يخف فيها ذلك.

• طريقة التسجيل العيوي : تستخدم هذه الطريقة

خاصة في الدراسات الديموغرافية، حيث يتم إعداد مكاتب خاصة، يتم التسجيل فيها مباشرة، أين يتقدم الأشخاص الى هذه المكاتب، و يتم تسجيل كل ما يطرأ على حالاتهم، من مواليد أو وفيات أو أمراض... الخ، وهذه الطريقة تكون منظمة في الغالب من طرف هيئات رسمية، و تلزم العقوبات في حالة عدم التقيد بها أو تفقد الإمتياز الذي قد تجرى لأجله. من أهم مزاياها أنها تضمن درجة عالية من الدقة في المعلومات، كما توفر المعلومات بسرعة وفي الوقت الذي يريده الباحث.

هذه أهم الطرق التي يمكن أن تستخدم في جمع البيانات الإحصائية، وكما يمكن إستخدام كل طريقة بإستقلال تام عن الأخرى، فإنه يمكن أيضا أن يتم إستخدام طريقتين أو أكثر في آن واحد لجمع البيانات الإحصائية.

د - الاستمارة الإحصائية : وهي الوثيقة التي يجري

من خلالها الحصول على المعلومات الإحصائية من المبحوثين

(أي عناصر المجتمع الإحصائي)، وهي من الأدوات الجوهرية في إنجاح العملية الإحصائية، ولكي تتم عملية جمع البيانات بسهولة و بأقل ما يمكن من الأخطاء وبأسرع وقت ممكن، يجب أن تتصف الإستمارة بما يلي:

• **من حيث الشكل:** يجب أن تكون ورقة الاستمارة من النوع الجيد الذي يتحمل الإستخدام الكثير، والذي يكون وقع القلم عليه جيدا، كما أن لون الإستمارة يجب أن يكون مقبولا ويفضل في ذلك اللون الأبيض كما يستحسن تلوين بعض مناطق الاستمارة أو تأطيرها بغية التحسيس على أهمية الأسئلة التي تتضمنها، كما يجب أن يكون حجم الاستمارة مناسباً، وأن لا يكثر من عدد الصفحات، حتى لا يؤثر ذلك على القوائم بعملية الإحصاء وعلى المستوجب، اذا كان تسجيل الإجابات ذاتيا. ومن الضروري أن تكون الأسئلة داخل الاستمارة مرتبة ترتيبا منطقيا، ومرقمة حسب الترتيب. وأخيرا من الضروري ترك مسافات كافية للإجابة أمام أو تحت كل سؤال.

• **من حيث المضمون:** ونقصد بذلك كيفية صياغة الأسئلة، وفي هذا الإطار يوصي بما يلي:

- أن تكون الأسئلة مختصرة قدر الإمكان وأن تتطلب إجابات قصيرة، وأن لا تطرح إلا الأسئلة الضرورية للدراسة الإحصائية.

- يجب ان تكون الأسئلة بسيطة ولا تتطلب إجراء بعض العمليات الحسابية، و أن تكون باللغة التي يفهمها معظم عناصر المجتمع الاحصائي.

- يجب ان تتحاشى الأسئلة الشخصية والعاطفية والأسئلة المخرجة بصفة عامة.

- أن تخدم وحدات قياس الأسئلة المتعلقة بالكميات.

- أن يشار الى المختصرات على الهامش.

و أخيرا من البديهي أن تحتوي الاستمارة في أعلاها، على مصدرها، ورقمها ونوع البحث الإحصائي وعنوانه والإشارة الى ضمان سرية المعلومات المحصل عليها من المبحوثين وعدم رسميتها والاستدلال بها في المتابعات القضائية أو الادارية التي قد يتخوف منها المستجوب.

١- أخطاء جمع البيانات الإحصائية : يمكن أن

ترتكب عدة أنواع من الأخطاء عند عملية جمع البيانات، نورد منها مايلي :

• **أخطاء المبحوثين:** وتكون إما متعمدة، ناتجة عن الخوف من المسئوليات التي قد تسترتب على الإجابات الصحيحة، كالخوف من فرض الضرائب أو المتابعة القضائية... الخ، أو الخوف من بعض الأوهام المنتشرة في المجتمع الإحصائي كالعين والحسد والطمع، أو تكون ناتجة عن عدم الوعي الإحصائي أو نتيجة لشخصية المستجوب التي قد تكون مهابة أو هزلية.

وقد تكون هذه الأخطاء غير متعمدة ناتجة عن عدم فهم السؤال أو الاسئلة المطروحة كأن يطرح السؤال بلغة غير مفهومة، أو عدم سماع السؤال جيدا، وقد تكون ناتجة عن سوء تصميم الاستمارة، أو عدم ملاءمة وقت الاستقصاء.

• **أخطاء الباحثين :** وهي الأخطاء التي يرتكبها

المستجوب عند تسجيله الإجابات، كأن يخطيء في تسجيل بعض الإجابات، أو يسجل إجابة في محل إجابة أخرى، أو أن ينسى طرح بعض الأسئلة... الخ.

٢- مراجعة البيانات : على الرغم من الاحتياطات

المتخذة سواء على مستوى المستجوبين أو على مستوى

شكليات الإستثمار فإنه يمكن أن ترتكب أخطاء عند جمع البيانات الإحصائية، لذلك فإنه يتم مراجعة الإستثمارات بعد الإنتهاء من جمع البيانات الإحصائية. وقبل الشروع في تبويب البيانات ينبغي القيام بما يصطلح عليه تدقيق أو مراجعة الإستثمارات و ذلك بهدف :

- التأكد من دقة المعلومات المحصل عليها.
- التأكد من عدم تناقض البيانات المسجلة في الإستثمارات مع واقع الظاهرة.
- التأكد من تسجيل البيانات حسب المواصفات المقدمة للباحث.
- التأكد من الحصول على جميع البيانات المطلوبة.
- النظر في الملاحظات التي يكون الباحث قد دوّنها على هامش الإستثمار أو في المكان المخصص لذلك.
- و نشير الى أنه في الدراسات الإحصائية الضخمة ينبغي مراجعة الإستثمارات ميدانيا قبل تسليمها، وذلك بإعادة قراءة ما كتب على المستجوب نفسه، إضافة الى ذلك فإنه ينبغي على المسئول على قاعات الفرز عند تسليم الإستثمارات التأكد من وجود إجابة على كل إستثمار وأن الإجابات مكتوبة بخط واضح وأنها مسجلة في الأماكن المخصصة لها، إضافة الى وجود إمضاء العداد و تاريخ إجراء التسجيل و عدد الإستثمارات المسلمة، وينبغي أن يجري ذلك في القاعات المعدة خصيصا عشية إجراء العملية.

بعد مراجعة الإستثمار يتم إتخاذ القرار إما بقبولها و اعتمادها في العملية الإحصائية أو رفضها إذا كانت تحتوي على أخطاء لا يمكن تصحيحها ويكون لها تأثير سلبي على نتائج العملية. أما إذا كان بالإمكان تصحيحها فإنه يتم ذلك دون اللجوء الى التشطيب أو مسح البيانات الخاطئة، و ينبغي لأجل

ذلك استخدام قلم من لون مغاير، ويستحسن في ذلك اللون الأحمر لكونه ملفتاً للإنتباه وباعتباره اللون المستخدم تقليدياً في التصحيح أو إبداء الملاحظات، كما ينبغي تدوين الملاحظات المتعلقة بالمراجعة و التصحيح، وذلك ليتمكن المراجع الموالي من إدراك ذلك في حالة وجود أكثر من مراجع واحد.

3 - المرحلة الثالثة: تبويب و عرض البيانات : بعد جمع البيانات الإحصائية في الاستثمارات (قد يكون عددها كبير جداً) فإن أكوام الاستثمارات تلك، لا يمكن بأي حال من الأحوال أن تعطينا نظرة وافية عن نتيجة العملية، لذلك لابد من اللجوء الى تصنيف و تبويب تلك البيانات، اي ان نجعلها في مجموعات متجانسة، تشترك في صفة أو خاصية واحدة أو عدة خواص، أي ان نجعلها في شكل فئات إحصائية، مقدمة في جداول مناسبة (انظر الفصل الموالي) ولتصنيف و تبويب البيانات فنيات خاصة، يلجأ اليها الإحصائي، حتى تصبح المعلومات عملية، ويمكن استخدامها من طرف الباحث نفسه، أو من طرف غيره، بعد تقديمها في نشرات خاصة أو دوريات عامة، وتصنيف و تبويب البيانات يجري حسب طبيعة البيانات ذاتها، وهو الشيء الذي سنبرزه من خلال الفصل الموالي.

4 - المرحلة الرابعة: تحليل البيانات واستقراء النتائج : تحليل البيانات هو وسيلة الحصول على الإجابات المطلوبة في إشكالية البحث الإحصائي، اذ تعتبر المراحل المشار اليها آنفاً ضرورية، حتى يتمكن الباحث من التحليل الإحصائي لجوانب الظاهرة المدروسة، ويتم ذلك عن طريق أدوات إحصائية كثيرة، منها البسيط ومنها المعقد، تسمح باستقراء النتائج واستخلاص مدلولها، الذي هو هدف البحث الإحصائي، وسنستعرض بعضها ابتداء من الفصل الرابع.

إن الحرص على اتباع خطوات المراحل الأربعة المشار إليها يجنب الباحث الكثير من المتاعب و المشاق التي يمكن أن يصادفها عمليا فيما لو لم يتقيد بها، إضافة الى تبذير الكثير من الأموال و الوقت دون فائدة، لأنه دون أدنى شك سوف يحصل على معلومات خاطئة قد تفيد الباحث فقط في جزء من الدراسة الإحصائية وقد لاتفيده بتاتا، والأخطر من ذلك أن نتائج الدراسة المتوصل إليها تكون خاطئة ما دامت المعلومات الأولية المحصل عليها خاطئة، الشيء الذي ينجر عنه قرارات خاطئة قد تعود بالسلب على الظاهرة المدروسة.

أسئلة و تمارين.

تمرين 1: اعط تعاريف للمصطلحات التالية:

1- تعداد.	5- صفة نوعية.
2- إحصائيات.	6- وحدة إحصائية.
3- علم الإحصاء.	7- مجتمع إحصائي.
4- بحث إحصائي.	8- عينة إحصائية.

تمرين 2 : وضح لماذا نضيف صفة العلمية على الإحصاء ؟.

تمرين 3: ماهي المواضيع التي يتناولها الإحصاء الوصفي ؟. حدد الفرق بين الإحصاء الوصفي والإحصاء الرياضي.

تمرين 4: في دراسة شاملة حول ظروف الطالب الجامعي، حدد الصفات النوعية والصفات الكمية المستقصى عنها.

تمرين 5: ما هي أهم المؤسسات الوطنية التي تقوم بجمع البيانات الإحصائية عن الظواهر الاقتصادية والاجتماعية ودراستها ؟.

تمرين 6: هل توجد مصلحة خاصة بالإحصائيات في المؤسسة التي تنتمي إليها ؟ حدد أهم الأدوار و الأعمال التي تقوم بها أو قامت بها.

تمرين 7: حدد مجالات إستخدام الإحصاء في الحياة اليومية، ووضح علاقة الإحصاء بـ : أ- العلوم الاقتصادية. ب- العلوم الطبية. ج- العلوم السياسية، مع إعطاء أمثلة كافية.

تمرين 8: وضح لماذا يلجأ الى إستخدام أسلوب العينات في بعض الدراسات الإحصائية؟.

تمرين 9: قم بإعداد إستمارة إحصائية، لحصر النفقات اليومية لطلبة الفرع الذي تنتمي إليه.

تمرين 10: وضع كيف يتم إختيار كل من: العينة الطبقية، العينة المنتظمة، العينة العنقودية.

تمرين 11: وضع كيف يتم إختيار عينة عشوائية تتكون من 95 عنصر من ضمن مجتمع إحصائي يتكون من 900 عنصر.

تمرين 12: قم بإعداد مخطط عمل شامل ومفصل، لإجراء دراسة إحصائية حول مستوى التغذية لسكان الولاية التي تقطن بها، وقدر تكلفة إنجاز المخطط بالأسعار الجارية.

تمرين 13: حدد أهم أخطاء جمع البيانات الإحصائية التي يمكن الوقوع فيها، عند إجراء البحث الإحصائي المطلوب في التمرين 12، وبين كيف يجب تفاديها.

تمرين 14 : ما هي الطرق والأساليب التي ينبغي إستخدامها في جمع البيانات الإحصائية حول المواضيع التالية :

- دراسة الحالة الإجتماعية لطلبة المعهد الذي تنتمي اليه.
- دراسة الإستهلاك الغذائي في الولاية التي تسكن فيها.
- دراسة مدى الإستجابة لدعوة تلقيح الأطفال ضد مرض معين.
- دراسة أعمار الأشخاص المستفيدين من إعانات إجتماعية تقدمها البلدية التي تسكن فيها.
- إجراء دراسة إحصائية عامة حول السكان والسكن في دولة

الفصل الثاني

تبويب البيانات الإحصائية.

كما سبقت الإشارة في منهجية البحث الإحصائي، فإن عملية تبويب البيانات الإحصائية، من الأعمال الضرورية التي لا بد منها حتى تصبح البيانات عملية، يمكن إستخدام أدوات التحليل الإحصائي لتفسيرها واستخلاص النتائج منها، ويتم تبويب البيانات عن طريق الجداول الإحصائية، إما يدوياً أو آلياً. وسنستعرض في هذا الفصل طريقة تبويب البيانات إضافة الى التعرض الى مختلف جوانب جداول التوزيعات التكرارية.

أولاً: مواصفات الجداول الإحصائية: إن تصنيف وتبويب البيانات الإحصائية يهدف الى وضع تلك البيانات في شكل مبسط، بحيث يمكن للباحث دراستها وتحليلها واستخلاص النتائج بكل سهولة، ويتم ذلك في جداول خاصة، تسمى بالجدول الإحصائية.

إن كل جدول إحصائي يتطلب مجموعة من المواصفات منها ما يلي:

- يجب أن يكون الجدول مصمماً بشكل مقبول، ويحمل كل عناصر الظاهرة المدروسة.

- يجب أن تكتب الأرقام بانتظام، بحيث توضع الآحاد تحت الآحاد والعشرات تحت العشرات و المئات تحت المئات... الخ، كما يجب أن يشار الى الأرقام غير المتوفرة أو غير الرسمية... بإشارات خاصة متعارف عليها منها ما يلي:

.... تعني أرقام غير متوفرة.

— تعني أرقام غير موجودة أصلاً.

x أرقام غير رسمية.

[] أرقام غير محسوبة ضمن المجموع.

- أن يكون للجدول عنوانا مختصرا قدر الإمكان، ومتضمنا لمحتوى الجدول.
- أن يحدد تحت العنوان أو ضمنه مكان الظاهرة وزمان حصرها.
- أن يشار الى وحدات قياس الظاهرة المدروسة، سواء في خانات الجدول أو على الهامش، أو تحت العنوان اذا كانت هناك وحدة قياس واحدة مشتركة لجميع عناصر الجدول.
- أن تحدد المختصرات التي قد تستخدم على هامش الجدول.
- في حالة استخدام أكثر من جدول لابد من ترقيم كل جدول ترقيما متتاليا حسب الترتيب، كما تجب الإشارة الى مصدر الأرقام.

و يتم تقديم البيانات في غالب الأحيان بشكل تكراري، ويسمى عرض البيانات بهذه الطريقة بالتوزيع التكراري، كما تسمى الجداول التي تعرض عن طريقها بالجداول التكرارية. والجداول التكرارية على أنواع، حسب طبيعة البيانات الإحصائية، منها الكمية و النوعية، البسيطة و المزدوجة، ومنها المتصلة و غير المتصلة، وذات الصفات البسيطة و المزدوجة، وسوف نستعرض في هذا البند كيفية افراغ المعلومات الأولية في جداول إحصائية بالطريقة اليدوية كما نستعرض كذلك، الأنواع المختلفة للتوزيعات التكرارية ذات الاستخدام الواسع. وقبل ذلك لا بأس أن نذكر بعناصر الوحدة الإحصائية التي قد تكون ذات صفات نوعية أو كمية و التي على أساسها يتم التصنيف، و تسمى هذه الصفات بالمتغيرات، ففي المتغيرات الكمية يمكن أن نصادف حالتين، هما:

- المتغيرات المتقطعة : وفي هذه الحالة تأخذ المتغيرة رقما واحدا محددًا كأن نقول : عدد العائلات الذين لديهم 4 أطفال في حي ما هو 10، فعدد الأطفال هنا محدد في رقم واحد فقط وليس في مجال.

- المتغيرات المستمرة : وفي هذه الحالة تأخذ الصفة الكمية أية قيمة ضمن مجال محدد، وذلك كالعمر مثلاً، كأن نقول عدد الأطفال الذين تتراوح أعمارهم بين 8 و 10 سنوات في قسم دراسي ما هو: 8 فالتغيرة هنا يمكن أن تأخذ أية قيمة ضمن هذا المجال مهما كانت بفواصلها، و بمعنى آخر أنها غير متقطعة أو مستمرة ضمنه.

ثانياً- الجداول التكرارية ذات الصفات النوعية : و هي الجداول التي تتضمن تكرارات صفات نوعية معينة للظاهرة المدروسة، كعدد المتزوجين، أو عدد حاملي شهادة الليسانس في تخصص ما، أو عدد العاطلين عن العمل، مثلاً، وقد تكون إما جداول تكرارية بسيطة أو جداول تكرارية مزدوجة (مركبة).

1- الجدول التكرارية البسيطة: وهي تحتوي على صفة نوعية واحدة فقط، ويتم افراغ البيانات فيها كما يلي :

- يرسم جدول مسودة يناسب حجم البيانات مقسم الى ثلاثة أعمدة.

- يوضع في العمود الأول الصفة، وفي الثاني وهو أكبر نسبياً التعداد، وفي العمود الثالث التكرار.

- يملأ الجدول حسب منهجية المثال التالي :

مثال 2-1: أخذت عينة عشوائية من الطلبة تتكون من 25 طالباً، ليتم إستقصاءهم عن شعب البكالوريا التي يحملونها، و تم ذلك من خلال ملأ استمارات خاصة، فكانت الاجابات في الإستمارات كما يلي :

أداب	أداب	علوم	أداب	رياضيات
أداب	علوم	علوم	رياضيات	علوم
علوم	رياضيات	علوم	علوم	رياضيات
علوم	علوم	رياضيات	رياضيات	أداب
علوم	رياضيات	علوم	أداب	علوم

المطلوب : أفرغ هذه البيانات في جدول مناسب.
نقوم بإعداد جدول حسب الشكل أدناه (جدول 1-2)،
وحتى نتجنب الخطأ خاصة إذا كان عدد الإستمارات أو عدد
البيانات كبيراً، نقوم بأخذ استمارة بعد استمارة، ونضع
تشطية عمودية صغيرة أمام الصفة التي تحتويها الاستمارة،
وذلك في عمود التعداد، وعندما نصل الى التشطية الخامسة
نضعها مقاطعة للأربعة الأولى، بحيث تشكل لنا زمرة تكون
من خمس تشطيات، ونستمر هكذا حتى ننتهي من تسجيل
كل الإستمارات.

ومن البديهي أن استخدام الزمر الخماسية على هذا المنوال،
يسهل لنا عملية الجمع عند الانتهاء من إفراغ البيانات في عمود
التعداد، وذلك ما يوضحه الجدول الموالي :

الفرع	التعداد	التكرار
أدب		6
علوم		12
رياضيات		7
المجموع		25

جدول 1-2

يعتبر الجدول 1-2 جدولاً أولياً، أي جدول مسودة، ويسمى
بجدول التفريغ، وهو يساعد على الإحصاء الدقيق، و يلاحظ
أن الزمر الخماسية في عمود التعداد تسهل لنا الحساب في
النهاية، ويتم حذف العمود الثاني، للحصول على الجدول
النهائي المطلوب، وذلك ما يظهره الجدول 2-2 أدناه.

توزيع عينة من الطلبة حسب فرع الدراسة.

الفرع	التكرار
أداب	6
علوم	12
رياضيات	7
المجموع	25

جدول 2-2

2- الجداول التكرارية المزدوجة للصفات النوعية:

وهي التي يتم التصنيف فيها على أساس صفتين نوعيتين في الوحدة الإحصائية، ولبيان ذلك نأخذ المثال التالي:

مثال 2-2: لمعرفة مدى تتبع طلبة الثانويات لنشرات الأخبار المتلفزة حسب الجنس، أختيرت عينة عشوائية من إحدى الثانويات تتكون من 10 طلبة، فكانت الإجابات في الإستمارات المعدة خصيصا لذلك كما يلي :

الرقم	الجنس	الرد
1	ذ	نعم
2	ذ	نعم
3	ذ	نعم
4	ث	لا
5	ذ	لا
الرقم	الجنس	الرد
6	ث	نعم
7	ث	لا
8	ذ	نعم
9	ذ	لا
10	ذ	نعم

ملاحظة : ذ : ذكر. ث : أنثى. نعم : يشاهد أو تشاهد. لا : لا يشاهد أو لا تشاهد.

المطلوب : افرغ البيانات أعلاه في جدول مناسب.

لإنجاز ذلك نقوم بإعداد جدول كما هو واضح أدناه، ونتبع الأسلوب المشار إليه سابقا في عملية التبويب، أي استخدام التشطيب العمودي، فنحصل على الجدول المسودة التالي :

الرد	لا يشاهد أو لا تشاهد		يشاهد أو تشاهد		الجنس
	تعداد	تكرار	تعداد	تكرار	
ذكور	11	2	111	5	7
إناث	11	2	1	1	3
المجموع		4		6	10

جدول 2-3

وباستبعاد التشطيطات العمودية، نحصل على الجدول المزدوج التالي:
جدول يظهر مدى مشاهدة عينة من الطلبة للأخبار
المنلفزة حسب الجنس.

الجنس	الرد	لا يشاهد أو لا تشاهد	يشاهد أو تشاهد	المجموع
ذكور		2	5	7
إناث		2	1	3
المجموع		4	6	10

جدول 2-4

يلاحظ إذن أن هذا الجدول، هو جدول تكراري مزدوج ذي صفات نوعية، أي مبني على أساس أكثر من خاصية (صفة) واحدة، إذ أن الصفة النوعية الأولى هي الجنس، والصفة النوعية الثانية هي مدى المشاهدة، وهو يعرض تلك البيانات التي كانت في الاستثمارات في شكل فوضوي، ويلخصها بوضوح تام، وهذا هو الهدف الأساسي من تبويب البيانات.

و ما تجدر الإشارة إليه هو أنه عند إعداد الجداول التكرارية ذات الصفات النوعية، فإنه يجب ترتيب صفات الظاهرة حسب أهميتها.

ثالثاً: الجداول التكرارية ذات الصفات الكمية : وهي الجداول التي يكون فيها التصنيف على أساس صفة كمية في الظاهرة، ومعنى آخر هي التي تعرض تكرارات كميات

الظواهر، كتوزيع العمال حسب الأجور المدفوعة لهم، أو حسب أعمارهم، وقد تكون هذه الجداول إما مستمرة (متصلة) أو غير مستمرة (منفصلة)، وفي كلتا الحالتين فإنه عند إفراغ البيانات الأولية في هذه الجداول، لابد من مراعات ترتيب القيم إما تصاعدياً أو تنازلياً، كما لابد من استخدام قاعدة العد الأساسية، وهي التشطيبات العمودية، وذلك لتفادي الأخطاء خاصة إذا ما كان عدد القيم كبيراً.

1- الجداول التكرارية غير المستمرة : وهي الجداول التي تظهر عدد تكرارات كمية واحدة محددة وممثلة في رقم واحد فقط، تسمى هذه الكمية بالفئة، وبمعنى آخر هي التي تكون فيها الصفة الكمية عبارة عن متغيرة متقطعة كما هي معرفة آنفاً. (أنظر المتغيرات المتقطعة).

مثال 2-3: البيانات التالية تعرض توزيع عمال مؤسسة ما، حسب عدد الأيام التي إشتغلوها في شهر جانفي.

عدد الأيام (الفئة)	عدد العمال (التكرار)
10	5
15	6
18	4
20	3
22	2
المجموع	20

جدول 2-5

هذه البيانات معروضة في جدول تكراري غير مستمر، حيث أن كل مجموعة من العمال إشتغلت عدداً محدداً في رقم واحد من الأيام، فعلى سبيل المثال عدد العمال الذين إشتغلوا 10 أيام تماماً هو 5، وعدد العمال الذين إشتغلوا 20 يوماً بالتمام هو 3، وهكذا. عدد الأيام يسمى بالفئة، وهو محدد في رقم واحد كما سبقت الإشارة، أي هو غير محصور ضمن مجال، و بالتالي

نقول أن طول الفئة (طول مجال الفئة)، معدوم، ونشير لذلك
إبتداءً من الآن بـ: $L=0$.

وتسمى مثل هذه الجداول بالجداول الكمية البسيطة غير
المستمرة، وهناك أيضاً الجداول الكمية المزدوجة غير
المستمرة، وهي المؤسسة بناء على كمتين في الظاهرة.

و ما تجدر الإشارة إليه، هو أنه عند القيام بعملية التبويب
اليدوي للبيانات في مثل هذه الجداول، فإننا نتبع نفس الطريقة
التي اتبعت في حالة تبويب البيانات ذات الصفت النوعية.

مثال 2-4: قامت مديرية الدراسات بأحد المعاهد بجمع بيانات عن عدد
أفراد الأسرة لعينة من الطلبة تتكون من 20 طالباً، من خلال استمارات
أعدت لهذا الغرض، فكانت النتائج الظاهرة في الإستمارة كما يلي :

6	7	3	2	10	2	8	6	6	5
7	5	9	7	6	3	4	10	7	7

المطلوب : تبويب هذه البيانات في جدول تكراري.

للقيام بما هو مطلوب، نقوم بإعداد جدول ثم نملأه، كما هو واضح أدناه:

عدد أفراد الأسرة (الفئة)	التعداد	عدد الطلبة (التكرار)
2		2
3		2
4		1
5		2
6		4
7		5
8		1
9		1
10		2
المجموع		20

جدول 2-6

باستبعاد عمود التعداد يصبح الجدول المطلوب كما يلي :

توزيع عينة من الطلبة حسب عدد أفراد الأسرة.

عدد أفراد الأسرة (الفئة)	عدد الطلبة (التكرار)
2	2
3	2
4	1
5	2
6	4
7	5
8	1
9	1
10	2
المجموع	20

جدول 2-7

الجدول المحصل عليه هو أيضا جدول تكراري كمي بسيط غير مستمر، أو منفصل، وذلك لإنفصال الفئات بعضها عن بعض، وعدم ارتباطها، فكل فئة مستقلة تماما عن الفئة التي تليها، وهي عبارة عن رقم واحد محدد تماما.

يرمز لقيمة الفئة بـ: x_i و لتكراراتها بـ: f_i ، حيث i : رقم الفئة.

2- الجداول التكرارية الكمية المستمرة (المتصلة):

وهي الجداول التي تكون فيها الظاهرة محصورة في مجال، بحيث يمكن أن تأخذ أية قيمة ضمنه، كأن نقول مثلا أن عدد الأطفال الذين تتراوح أطوالهم بين 100 سم و 120 سم في قسم دراسي ما هو 10 أطفال، فالظاهرة في هذا المثال وهي أطوال الأطفال يمكن أن تأخذ أية قيمة ضمن المجال 100 و 120 سم، و بمعنى آخر هي الجداول التي تكون فيها الصفة الكمية عبارة عن متغيرة مستمرة كما هي معرفة سابقا، و يتم استخدام هذه الطريقة في عرض البيانات إذا كان عددها كبيرا، وذلك لتقليصها، إذ أن هدف التبويب هو عرض البيانات بأقل حيز ممكن وبأقصى وضوح، فيتم حينئذ تحديد فئات طولها أكبر من الصفر، ويتم إهمال القيم التي تقع داخل مجال الفئات.

يسمى طول مجال الفئة بمدى الفئة أو طول الفئة، ونرمز له بالحرف L

مثال 2-5: البيانات التالية خاصة بالنفقات اليومية لعينة من الطلبة بالدينار، والمطلوب تبويبها في جدول تكراري يجعل طول الفئة : $L = 4$ دينار.

13	8	5	17	20	7	6	15	4
11	23	15	12	5	25	9	7	4
22	9	9	22	18	7	8	17	6
4	21	16	20	19	12	7	8	4

لتبويب هذه البيانات نقوم بتصميم جدول تكراري، ثم نحدد أدنى قيمة ضمن مجموعة البيانات لتكون الحد الأدنى لأول فئة، وهي في مثالنا 4، وبما أن طول الفئات المطلوب أن نبوب على أساسه البيانات هو : $L = 4$ ، لذلك فإن الحد الأعلى للفئة الأولى هو 8، ويكون الحد الأعلى للفئة الأولى هو الحد الأدنى للفئة الثانية، وهكذا، نحصل على الفئات كما هي واضحة في الجدول 2-8 أدناه، ثم نعيد نفس مبدأ العد المشار إليه آنفاً، وهو استخدام التشطيطات العمودية، وذلك لتفادي الأخطاء، خاصة عندما يكون عدد البيانات كبيراً جداً، ويكون الجدول الموالي هو الجدول المطلوب رقم 2-8.

توزيع عينة من الطلبة حسب النفقات اليومية بالدينار.

رقم الفئة	الفئة	التكرار
1	8-4	12
2	12-8	7
3	16-12	5
4	20-16	5
5	24-20	6
6	28-24	1
مج		36

جدول 2-8

الفئة الأولى (8-4) مثلاً، تعني الطلبة الذين يتفوقون من 4 الى أقل من 8 دينار، عددهم 12 طالباً، فالعبارة (-)، تعني الى أقل

من. يسمى الطرف الأول للفئة بالحد الأدنى للفئة، أما طرفها الثاني فيسمى بالحد الأعلى للفئة، ومن الواضح في الجدول 2-8 أن طول الفئات متساو، ويساوي في كل فئة 4، و طول الفئة عبارة عن الفرق بين حدها الأعلى وحدها الأدنى، أي:

$$L_i = T_{i+1} - T_i \quad 1-2$$

حيث : L_i : طول الفئة i . T_{i+1} : الحد الأعلى للفئة i . T_i : الحد الأدنى للفئة i . من الواضح أن التوزيع المحصل عليه في الجدول 2-8، لا يتضمن جميع قيم المعلومات الأولية كما وردت أصلاً، لكنها متضمنة داخل الفئات، وتم إهمالها، لأجل حصر البيانات و تقديمها في صورة أكثر تلخيصاً.

عند تصميم أي جدول تكراري مستمر، تكون مشكلة الباحث الإحصائي هي، اختيار طول الفئة المناسب لعرض البيانات، وكذلك تحديد عدد الفئات المناسب لذلك، ويتم تحديد ذلك في غالب الأحيان، بصورة تضمن عدم فقدان البيانات لأهميتها في التحليل الإحصائي، وبالتالي يعتمد ذلك على ذاتية الإحصائي، أي إلى مدى مهارته وفنيته في التعامل مع البيانات، غير أن هناك قاعدة عامة، يتم بموجبها تحديد طول الفئة المناسب، لتبويب البيانات في جداول تكرارية مستمرة، تضمن عدم فقدان البيانات لأهميتها في التحليل الإحصائي وهي كما يلي :

*** تحديد طول الفئة:** إن تحديد طول الفئة يساعد على تحديد عدد الفئات وبالتالي حجم الجدول، إذ كلما كان طول الفئة كبيراً كلما كان حجم الجدول صغيراً، والعكس صحيح، و لتحديد طول الفئة يتم استخدام قاعدة ستيرجس (H.A.Sturges)، التي تعطى كما يلي :

$$L = \frac{W}{1 + 3.322 \log N} \quad 2-2$$

حيث : L : طول الفئة. N : عدد القيم. W : المدى، وهو الفرق بين أكبر قيمة ضمن مجموعة القيم، وأصغر قيمة ضمنها أي :

$$W = X_{\text{Max}} - X_{\text{Min}} \quad 3-2$$

حيث : X_{Max} : أعظم (أكبر) قيمة ضمن مجموعة القيم.
 X_{Min} : أدنى (أصغر) قيمة ضمن مجموعة القيم.

وكما سبقت الإشارة فإن هذه القاعدة تعطينا طول الفئات المناسب لإفراغ مجموعة البيانات في جدول تكراري مستمر، غير أن الإلتزام بها ليس إجبارياً، بل يبقى تحديد طول الفئة أمراً فنياً، يعود للإحصائي القائم بالعملية.

• **تحديد عدد الفئات** : يحدد عدد الفئات باستخدام القاعدة التالية :

$$N_c = \frac{W}{L} \quad 4-2$$

حيث N_c : عدد الفئات.

و من المعادلة 2-2 يمكن أن نكتب :

$$W = L(1 + 3.322 \log N) \quad 5-2$$

بتعويض 5-2 في 4-2، نجد أنه يمكن كتابة 4-2 أيضاً على النحو التالي :

$$N_c = 1 + 3.322 \log N \quad 6-2$$

مثال 2-6: أوجد طول الفئات المناسب لإفراغ البيانات التالية في جدول تكراري مستمر، ثم بوبها في جدول تكراري مستمر حسب طول الفئات المحصل عليه.

24	27	32	44	33	16	20	32	34	25
60	57	45	55	60	32	56	54	34	27
53	44	28	33	57	56	25	54	44	23
34	53	64	62	61	41	51	32	43	36
43	43	62	47	41	37	52	54	42	55

الإجابة : أكبر قيمة في مجموعة البيانات هي : 64، بينما أصغر قيمة هي: 16، و عدد القيم هو : 50، و عليه فبتطبيق المعادلة 2-2 نجد:

$$L = \frac{64 - 16}{1 + 3.322 \log 50} \cong 7$$

إذن طول الفئات المناسب لإفراغ هذه البيانات في جدول تكراري مستمر هو: 7.

أما لإيجاد عدد الفئات المناسب فيتم استخدام إما المعادلة 2-4 أو 2-6، وكلتاهما، تعطيان تقريبا نفس النتيجة و هي 7، وبالتالي يكون الجدول المطلوب هو:

i	الفئة	f _i
1	23-16	2
2	30-23	7
3	37-30	10
4	44-37	7
5	51-44	5
6	58-51	13
7	65-58	6
مج		50

جدول 2-9

كما سبقت الإشارة فإنه بالرغم من هذه القواعد لحساب كل من طول الفئات وعددها المناسب، لإفراغ البيانات في جدول تكراري مستمر، إلا أن ذلك يبقى أمرا فنيا يعود الى الباحث نفسه، ليتصرف حسب طبيعة البيانات من حيث عددها وكثافتها وحجمها، ولتسهيل الحسابات ينصح في الغالب بأن يكون طول الفئة من مضاعفات العدد 5 أو من مضاعفات عددا زوجيا، كما يفضل أن يحدد الحد الأدنى لأول فئة، بحيث

يجعل عملية حساب الحدود العليا و الدنيا للفئات الموالية سهلة، ففي مثالنا السابق، كان ينبغي أن يكون الحد الأدنى لأول فئة هو 14 بدلا من 16، حتى وان كانت القيمة 14 لا توجد ضمن المعلومات الأولية، حيث أن هذا الرقم من مضاعفات العدد 7، أي من مضاعفات طول الفئة، وبالتالي يسهل لنا عملية حساب حدود الفئات الموالية، حتى وإن أدى ذلك الى إضافة فئة عن العدد الذي حددته لنا قاعدة ستيرجس.

رابعاً: أنواع التوزيعات التكرارية المستمرة: تقدم الجداول التكرارية المستمرة بعدة صيغ منها ما يلي :

1- التوزيع التكراري المغلق : وفيه يكون الحد الأدنى لأول فئة والحد الأعلى لآخر فئة محددين، وقد يكون فيه مدى الفئات متساوياً، وفي هذه الحالة يسمى بالتوزيع التكراري المنتظم، وقد يكون مدى الفئات غير متساوياً، وفي هذه الحالة يسمى بالتوزيع التكراري غير المنتظم، ويلجأ اليه الباحث عندما تكون البيانات الإحصائية كبيرة التشتت، وكثيرة التمرکز في بعض الزمر .

2- التوزيع التكراري المفتوح : وفيه يكون إما الحد الأدنى لأول فئة غير محدد، أو الحد الأعلى لآخر فئة غير محدد، أو الحدين معاً . ففي الحالة الأولى يسمى بالتوزيع التكراري المفتوح من الأسفل، أما في الحالة الثانية، فيسمى بالتوزيع التكراري المفتوح من الأعلى، أما في الحالة الأخيرة فيسمى بالتوزيع التكراري المفتوح الطرفين .

أمثلة 2-7:

توزيع تكراري مغلق	توزيع تكراري مفتوح من أعلى	توزيع تكراري مفتوح من أسفل	توزيع تكراري مفتوح الطرفين
الفئات	الفئات	الفئات	الفئات
5	04 - 08	5	أقل من 8
7	12 - 08	7	12 - 08
3	16 - 12	3	16 - 12
1	20 - 16	1	16 فأكثر
جدول 2-10	جدول 2-12	جدول 2-11	جدول 2-13

3- التوزيعات التكرارية المتجمعة : إن التوزيعات التكرارية كما رأينا لحد الآن، تظهر لنا فقط عدد مرات تكرار الفئة، ولمعرفة عدد التكرارات التي تقل، أو تساوي وتزيد عن حد معين من حدود الفئات بسهولة، فإنه يتم اللجوء إلى جداول التوزيعات التكرارية المتجمعة، وهي التوزيع التكراري المتجمع النازل، و التوزيع التكراري المتجمع الصاعد.

* التوزيع التكراري المتجمع الصاعد : يستخدم لفرض المعرفة السريعة لعدد أو نسبة التكرارات التي تقل عن حد معين من حدود الفئات، وفي حساب بعض مقاييس التزعة المركزية (أنظر: فصل 4)، في هذا التوزيع يكون عدد التكرارات التي تقل عن الحد الأعلى لأول فئة يساوي عدد تكرارات أول فئة، وعدد التكرارات التي تقل عن الحد الأعلى للفئة الثانية تساوي إلى عدد تكرارات الفئة الأولى والثانية، أما عدد التكرارات التي تقل عن الحد الأعلى للفئة الثالثة فيساوي إلى مجموع تكرارات الفئة الأولى والثانية والثالثة، وهكذا، يستمر التجميع حتى الوصول إلى التكرارات التي تقل عن الحد الأعلى لآخر فئة، حيث يساوي إلى مجموع التكرارات.

مثال 2-8 : البيانات التالية تظهر عدد سكان دولة ما حسب

i	الفئة (سنة)	عدد السكان (10 ⁶ نسمة)
1	20-10	8
2	30-20	6
3	40-30	5
4	50-40	4
5	60-50	1
6	70-60	1
مجموع		25

جدول 2-14

المطلوب : أوجد التوزيع التكراري المتجمع الصاعد.

لو طرح علينا السؤال التالي : ماهو عدد السكان الذين تتراوح أعمارهم بين 30 وأقل من 40 سنة، لكنت الإجابة من الجدول السابق مباشرة وهي 5 مليون نسمة، غير أنه لو طرح علينا السؤال : ماهو عدد السكان الذين تقل أعمارهم عن 40 سنة ؟، لكنت الإجابة تتطلب وقتا لإجراء الحسابات قبل الإجابة على السؤال، اذ يتطلب الأمر جمع تكرارات الفئة الأولى والثانية و الثالثة والرابعة، أي $(5+6+8) = 19$ مليون نسمة، وعلى هذا المنوال يتم إيجاد التكرارات التي تقل عن أي حد من حدود الفئات، وهو ما يظهره الجدول 2-15 أدناه.

i	الفئة	f _i	التكرار المتجمع الصاعد
			الحد الأعلى
2	20-10	8	أقل من 20
3	30-20	6	أقل من 30
4	40-30	5	أقل من 40
5	50-40	4	أقل من 50
6	60-50	1	أقل من 60
7	70-60	1	أقل من 70
مج		25	

جدول 2-15

من الجدول السابق يمكن معرفة التكرارات التي تقل عن أي حد من حدود الفئات المحددة، ويلاحظ أن التجميع يجري بصفة تصاعدية، أي من الأدنى إلى الأعلى، لذلك سمي هذا التوزيع بالتوزيع التكراري المتجمع الصاعد، ويرمز للتكرارات المتجمعة الصاعدة بسهم إلى الأعلى.

والجدير بالذكر أنه في حالة التوزيعات التكرارية المفتوحة من الأعلى، أي الحد الأعلى لآخر فئة غير محدد، فيجب أن نكتب "مجموع التكرارات" بدل "أقل من"، وهذا عند الوصول لآخر فئة.

*** التوزيع التكراري المتجمع النازل:** يستخدم لغرض المعرفة السريعة لعدد التكرارات التي تساوي أو تزيد عن حد معين من حدود الفئات. في هذا التوزيع يكون عدد التكرارات التي تساوي أو تزيد عن الحد الأدنى لأول فئة مساويا، لمجموع التكرارات، أما التكرارات التي تساوي أو تزيد عن الحد الأدنى للفئة الثانية، فتكون مساوية لمجموع التكرارات منقوصا منها تكرار الفئة الأولى، وبالمثل تكون التكرارات التي تساوي أو تزيد عن الحد الأدنى للفئة الثالثة مساوية إلى مجموع التكرارات منقوصا منها تكرارات الفئة الأولى والثانية، وهكذا حتى نصل إلى التكرارات التي تساوي أو تزيد عن الحد الأدنى لآخر فئة حيث يساوي إلى تكرارات آخر فئة، وبمعنى آخر يكون مجموع التكرارات التي تساوي أو تزيد عن الحد الأدنى لآخر فئة يساوي إلى تكرار آخر فئة، أما مجموع التكرارات التي تساوي أو تزيد عن الحد الأدنى للفئة ما قبل الأخير فيساوي إلى مجموع تكرارات آخر فئة و الفئة التي تسبقها وهكذا...

مثال 2-9: أوجد التكرار المتجمع النازل لبيانات المثال 2-8.

بتطبيق فكرة التكرار المتجمع النازل، كما هي واضحة أعلاه
نحصل على الجدول المطلوب التالي :

i	الفئة	f_i	التكرار المتجمع النازل
			الحد الأدنى ✓
2	20-10	8	10 فاكتر 25
3	30-20	6	20 فاكتر 17
4	40-30	5	30 فاكتر 11
5	50-40	4	40 فاكتر 6
6	60-50	1	50 فاكتر 2
7	70-60	1	60 فاكتر 1
مج		25	

جدول 2-16

من الجدول 2-16 يمكن معرفة وبكل سهولة عدد التكرارات التي تساوي أو تزيد عن أي حد من حدود الفئات المتضمنة في البيانات الأولية، وفيه يكون التكرار الذي يساوي أو يزيد عن الحد الأدنى لآخر فئة مساويا الى تكرار آخر فئة، والتكرار الذي يساوي أو يزيد عن الحد الأدنى لأول فئة مساويا الى مجموع التكرارات . ويرمز للتكرارات المتجمعة النازلة بسهم الى الأسفل .

والجدير بالذكر أن هناك من يعتبر التكرار المتجمع الصاعد كما قدمناه نازلا، و التكرار المتجمع النازل كما قدمناه صاعدا، و ذلك لاختلاف وجهة النظر بالنسبة للتجميع.

4- التوزيع التكراري النسبي و المائوي : لمعرفة نسبة تكرار أية فئة من مجموع التكرارات، يتم إيجاد التكرار النسبي، وذلك بقسمة تكرار كل فئة على مجموع التكرارات، فالتكرار النسبي للفئة اذن هو نسبة تكرار تلك الفئة الى مجموع التكرارات، وغالبا ما يتم ضربه في 100 للحصول على ما

يسمى بالتكرار المائوي، وهو يعطي النسبة المئوية لتكرار كل فئة ضمن مجموع التكرارات، أي:

$$f_i \% = \frac{f_i}{\sum_{i=1}^n f_i} \times 100$$

7-2

حيث : $f_i\%$: التكرار المائوي للفئة i و f_i تكرار الفئة i .

مثال 2-10: الجدول التالي يظهر توزيع عينة من الطلبة حسب النتائج المحصل عليها في قسم دراسي يتكون من 10 طلبة، والمطلوب هو إيجاد التكرار المائوي .

عدد الطلبة	العلامة	i
1	5	1
4	6	2
2	7	3
1	8	4
1	9	5
1	10	6
10		مج

جدول 2-17

الإجابة : بتطبيق المعادلة رقم 2-7 نجد :

$$f_1 \% = \frac{1}{10} \times 100 = 10 \%$$

يعني هذا أن 10 % من الطلبة حصلوا على علامة تقدر بـ: 5، ويتم إيجاد بقية التكرارات المئوية بنفس الطريقة، وتظهر النتائج في الجدول التالي :

$f_i \%$	f_i	العلامة	i
10	1	5	1
40	4	6	2
20	2	7	3
10	1	8	4
10	1	9	5
10	1	10	6
100	10		مج

جدول 2-18

في غالب الأحيان يتم استخدام التكرار المائوي على النحو الذي رأينا آنفاً، غير أنه يتم أحياناً استخدام التكرار الألفي، أو النسبة الألفية بدل النسبة المئوية، وفي هذه الحالة يتم ضرب التكرار النسبي في 1000 بدل 100.

وما تجدر الإشارة إليه هو أن مجموع التكرارات النسبية يكون دائماً مساوياً إلى الواحد الصحيح، بينما مجموع التكرارات المئوية فيساوي إلى 100، كما هو واضح في المثال أعلاه.

تمارين.

تمرين 1: مصنع به 49 عاملاً، حصرت حالتهم العائلية من خلال إستمارات خاصة وزعت عليهم فكانت الإجابات على النحو التالي:

أعزب	أعزب	متزوج	أعزب	أعزب	أرمل	متزوج
أعزب	أعزب	متزوج	متزوج	متزوج	أعزب	مطلق
أعزب	أعزب	متزوج	مطلق	متزوج	أعزب	أعزب
مطلق	مطلق	متزوج	متزوج	مطلق	أعزب	أعزب
أعزب	متزوج	متزوج	مطلق	أعزب	متزوج	متزوج
مطلق	أعزب	أعزب	أعزب	متزوج	أعزب	أعزب
متزوج	أعزب	أعزب	أ مطلق	أ مطلق	متزوج	أعزب

المطلوب: 1- ما نوع هذه البيانات. 2- بوب هذه البيانات في جدول مناسب.

تمرين 2: لمعرفة شعب البكالوريا للطلبة المسجلين في الفوج الأول من السنة الأولى علوم إقتصادية حسب الجنس، أعدت إستمارات خاصة لأجل ذلك، وكانت الإجابات من خلالها كما يلي:

شعبة	جنس	شعبة	جنس	شعبة	جنس	شعبة	جنس	شعبة	جنس	شعبة	جنس
ع	ث	ر	ث	ر	ث	ر	ث	ر	ث	ع	ث
أ	ث	ر	ذ	ع	ث	ع	ث	ع	ث	ر	ذ
أ	ث	أ	ذ	ر	ذ	ع	ذ	ع	ذ	ع	ذ
ر	ث	ر	ذ	ع	أ	ث	ذ	ر	ذ	ع	ث
ر	ث	ر	ث	ع	ث	ع	ذ	أ	ذ	ع	ذ
ر	ذ	ع	ذ	ع	أ	ث	أ	ث	أ	ث	ذ

ملاحظة : أ : أداب. ر : رياضيات. ع : علوم. ذ : ذكر. ث : أنثى.

المطلوب: 1- إفرغ هذه البيانات في جدول مناسب. ما نوع هذا الجدول ؟.

تمرين 3: حصرت الأجور الأسبوعية لعمال أحد المصانع بمئات الدينارات، كما يلي :

41	40	44	39	38	35	36	36	37	35
42	41	43	40	44	41	40	43	44	42
42	47	48	49	45	41	43	41	42	40
48	52	55	46	50	45	48	41	45	48
47	41	58	54	59	50	55	51	49	43

المطلوب: 1- أوجد طول الفئات المناسب لإفراغ هذه البيانات في جدول تكراري مستمر. ما هو عدد الفئات عندئذ ؟ وضح كيف يتم حساب ذلك.

2- إذا كان $L=400$ دينار، إفراغ هذه البيانات في جدول تكراري

مستمر.

3- من البيانات المحصل عليها من السؤال: 1- أوجد التكرار المائوي.

ب- أوجد التكرارين المتجمعين الصاعد و النازل ثم التكرارين المائويين المتجمعين الصاعد و النازل.

تمرين 4: البيانات التالية خاصة بمساحة وعدد سكان الدول العربية سنة 2003 حسب البيانات الديمغرافية للأمم المتحدة.

(عدد السكان بالآلاف، المساحة بالكيلومتر مربع)

الدولة	عدد السكان	المساحة	الدولة	عدد السكان	المساحة
المغرب	30566	446550	الجزائر	31800	2381741
موريتانيا	2893	1025520	السعودية	24217	2149690
عمان	2851	212457	البحرين	724	678
قطر	610	11000	مصر	71931	1001449
الصومال	9890	637657	إ.ع. المتحدة	2995	83600
السودان	33610	2505813	العراق	25175	438317
سوريا	17800	185180	الأردن	5473	97740
تونس	9832	163610	الكويت	2521	17818
اليمن	20010	527968	لبنان	3653	10400
ح.ع. الصحراوية	-	266000	ليبيا	5551	1759540

المطلوب: 1- أوجد مساحة الوطن العربي، ومساحة المغرب العربي الكبير، والكثافة السكانية فيهما. **2-** أوجد النسبة المئوية لمساحة كل دولة من المساحة الكلية للوطن العربي. **3-** أوجد النسبة المئوية لسكان كل دولة من مجموع سكان الوطن العربي. **4-** ماهي النسبة المئوية لمساحة دول المغرب العربي الكبير من مجموع مساحة الوطن العربي. و ماهي النسبة المئوية لسكان المغرب العربي الكبير من مجموع سكان الوطن العربي. نفس السؤال بالنسبة لـ دول الخليج العربي.

تمرين 5: الجدولين التاليين يعرضان توزيع السكان المقيمين في الجزائر حسب فئات الأعمار سنتي 2000 و 1993 بالآلاف على التوالي.

توزيع سكان الجزائر حسب فئات الأعمار و الجنس سنة 2000

المصدر: الجزائر بالأرقام العدد 31 الديوان الوطني للإحصائيات / www.ons.dz

الفترة/سنة	ذكور	إناث	الفترة/سنة	ذكور	إناث
0-4	1534	1495	45-49	626	611
5-9	1799	1720	50-54	441	447
10-14	1933	1860	55-59	346	362
15-19	1890	1817	60-64	316	334
20-24	1617	1571	65-69	267	283
25-29	1352	1331	70-74	185	194
30-34	1149	1136	75-79	112	115
35-39	930	918	80 فأكثر	112	125
40-44	751	744	مجموع	15360	15026

توزيع سكان الجزائر حسب فئات الأعمار و الجنس سنة 1993

المصدر: المجموعة الإحصائية الديوان الوطني للإحصائيات 1994. ص: 22

الفترة/سنة	ذكور	إناث	الفترة/سنة	ذكور	إناث
0-4	1886	1804	40-44	564	542
5-9	1856	1776	45-49	410	418
10-14	1750	1677	50-54	335	357
15-19	1528	1461	55-59	304	329
20-24	1311	1268	60-64	254	274
25-29	1121	1101	65-69	191	204
30-34	919	892	70-74	136	145
35-39	731	694	75 فأكثر	175	183

1- هل هذين التوزيعين صحيحين إحصائيا، إذا كانا غير صحيحين وضح لماذا؟.

2- حول هذين التوزيعين الى توزيعين تكراريين مستمرين بجعل : $L = 5$ سنوات، بدل : $L = 4$ ، مع الإحتفاظ بنفس عدد السكان لكل فئة كما هي في الجدول.

3- من الجدول المحصل عليه من السؤال 2:

- أ- أوجد مجموع سكان كل فئة للتوزيعين
- ب- أوجد التكرار المائوي للجنسين لكل فئة (نسبة كل فئة) و لكل توزيع.
- ج- أوجد التكرارين المتجمعين الصاعد والنازل للجنسين و للتوزيعين.
- د- أوجد التكرارين المائويين المتجمعين الصاعد والنازل للجنسين و لكل توزيع.

4- أوجد مجموع سكان كل فئة للتوزيعين و أجب على الأسئلة التالية:

- أ- ماهي نسبة السكان الذين تقل أعمارهم عن : 10 سنوات، 20 سنة، 30 سنة، 40 سنة من كل توزيع.
 - ب- ما هي نسبة السكان الذين تساوي أو تزيد أعمارهم عن: 10 سنوات، 20 سنة، 30 سنة، 40 سنة من كل توزيع.
 - ج- إذا كان سن الشباب يتراوح بين 20 و 40 سنة، ما هي نسبة شباب الجزائر سنة 1993 و سنة 2000. ماذا تستنتج.
 - د- إذا كان سن الشيخوخة يبدأ من 60 سنة، فماهي نسبة الشيخوخة في الجزائر سنة 2000 و سنة 1993. ماذا تستنتج.
- 8- أعد تقديم البيانات عن طريق جدول تكراري مستمر بجعل طول الفئة: $L = 10$ للتوزيعين، ثم أعد الإجابة على كل فروع السؤال 3 .

الفصل الثالث

العرض البياني.

إن عرض البيانات الإحصائية بالطرق الإنشائية أو عن طريق الجداول الإحصائية كما رأينا، قد يكون مملاً أو صعب الفهم، خاصة إذا ما كانت هذه البيانات موجهة من خلال الجرائد أو المجلات أو لوحات التبليغ إلى عامة الناس، لذا يلجأ في الكثير من الأحيان إلى تقديم البيانات الإحصائية عن طريق الرسوم، بأساليب معينة، ووفق خصائص محددة، بغية تبسيطها وتسهيل تحليلها.

أولاً: مواصفات الأشكال البيانية: يجب أن تتصف الأشكال البيانية التي تعرض عن طريقها البيانات الإحصائية بمجموعة من المواصفات منها ما يلي:

- أن يكون الشكل جيداً من حيث التقديم وملفناً للإتباع.
- أن يكون له عنواناً في غاية الوضوح والاختصار، ومحدداً لكل ما يعبر عنه الرسم من معلومات بما فيها المكان والزمان، ووحدة القياس.
- أن تكون وحدات القياس محددة بدقة.
- أن يكون مقياس الرسم واضحاً، إذا كانت البيانات مقدمة على معلم.
- أن يشار إلى الألوان والرموز المستخدمة في تبيان محتويات الرسم، على جانب الرسم أو على الهامش لتوضيح معناها.
- أن يذكر مصدر الرسم أو المعلومات التي يعبر عنها، أسفل أو على الهامش.
- و هناك أشكال ورسومات كثيرة مستخدمة في عرض البيانات الإحصائية منها ما يلي :

ثانياً: الرسوم التكرارية : هي الرسوم التي تعرض البيانات المقدمة في جداول تكرارية، ومنها :

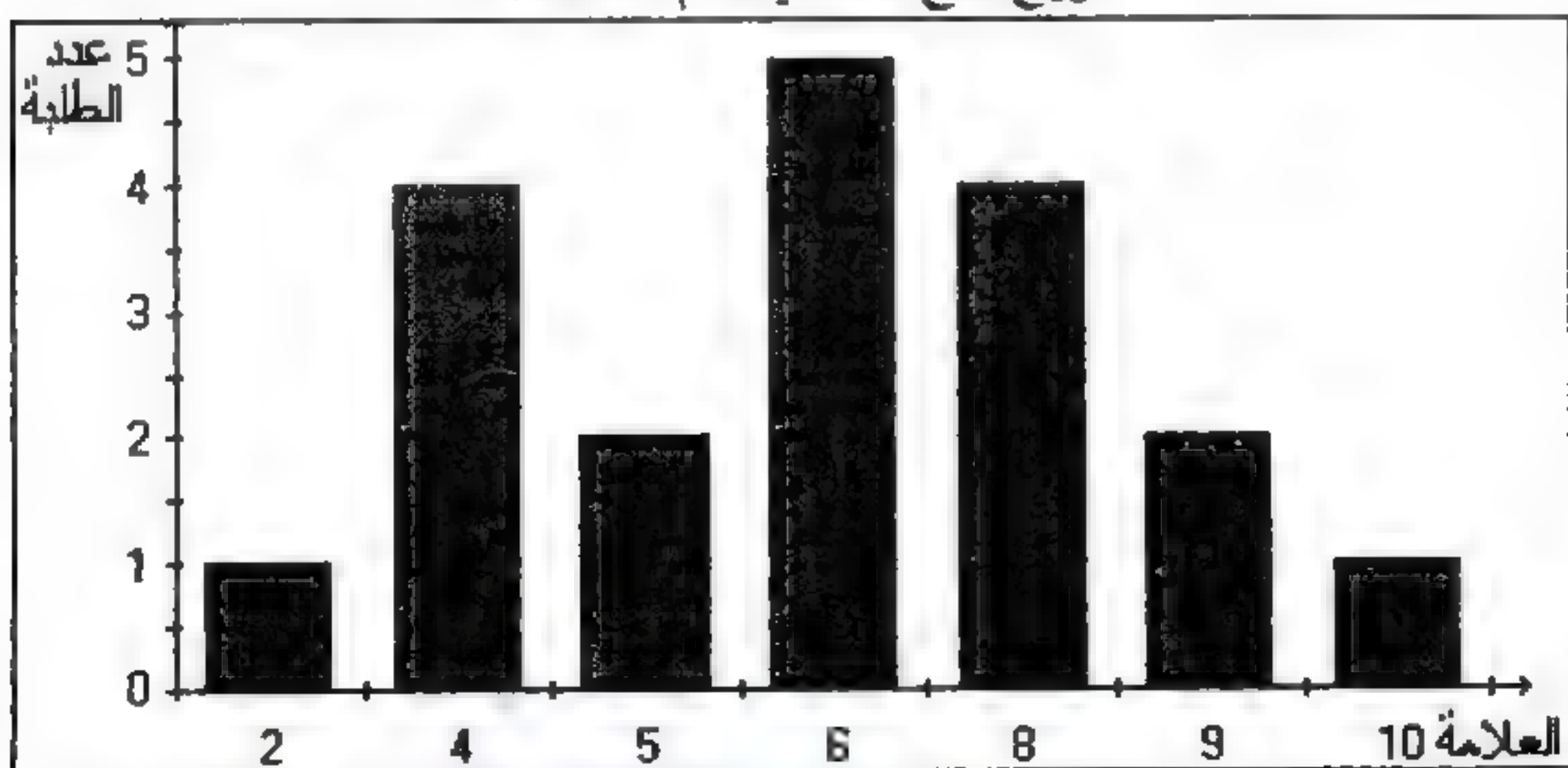
1- الأعمدة التكرارية : وهي تستخدم لتقديم البيانات الإحصائية التي لامتدى لفئاتها، و البيانات ذات الصفات النوعية، و لتقديم البيانات بهذه الطريقة، يتم إعداد معلم متعامد، بحيث توضع الفئات (سواء كانت ذات صفات كمية، أو ذات صفات نوعية) على المحور الأفقي، أما على المحور العمودي فيتم وضع التكرارات، و يتم الرسم كما في الشكل 1-3.

مثال 1-3: البيانات الواردة في الجدول التالي تظهر، توزيع عينة من الطلبة حسب النتائج المحصل عليها في قسم به 19 طالب، والمطلوب تقديمها على شكل أعمدة تكرارية.

العلامة	2	4	5	6	8	9	10	مج
عدد الطلبة	1	4	2	5	4	2	1	19

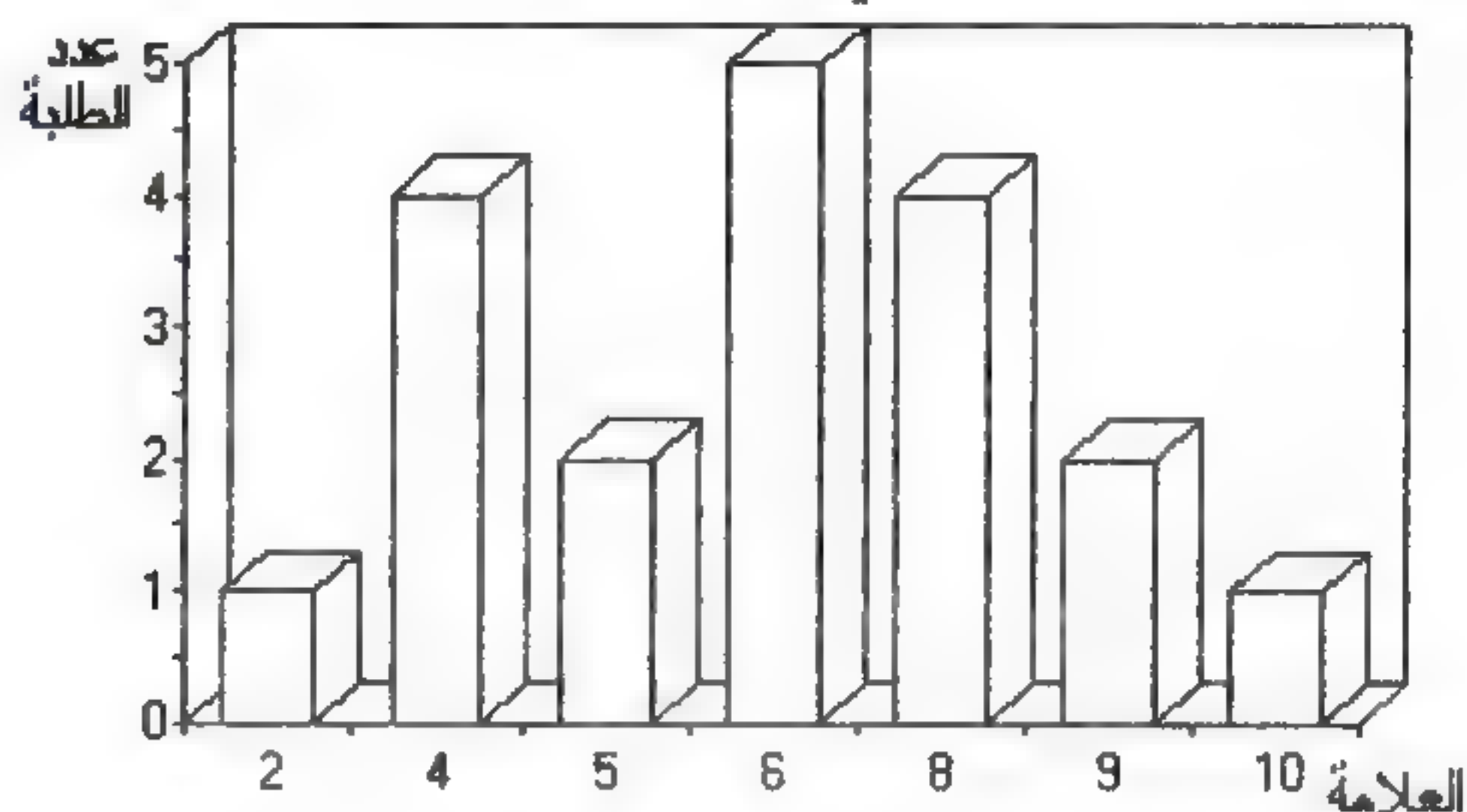
جدول 1-3

الإجابة: بتطبيق القاعدة أعلاه نحصل على البيان التالي:
توزيع نتائج الطلبة في قسم به 19 طالباً.



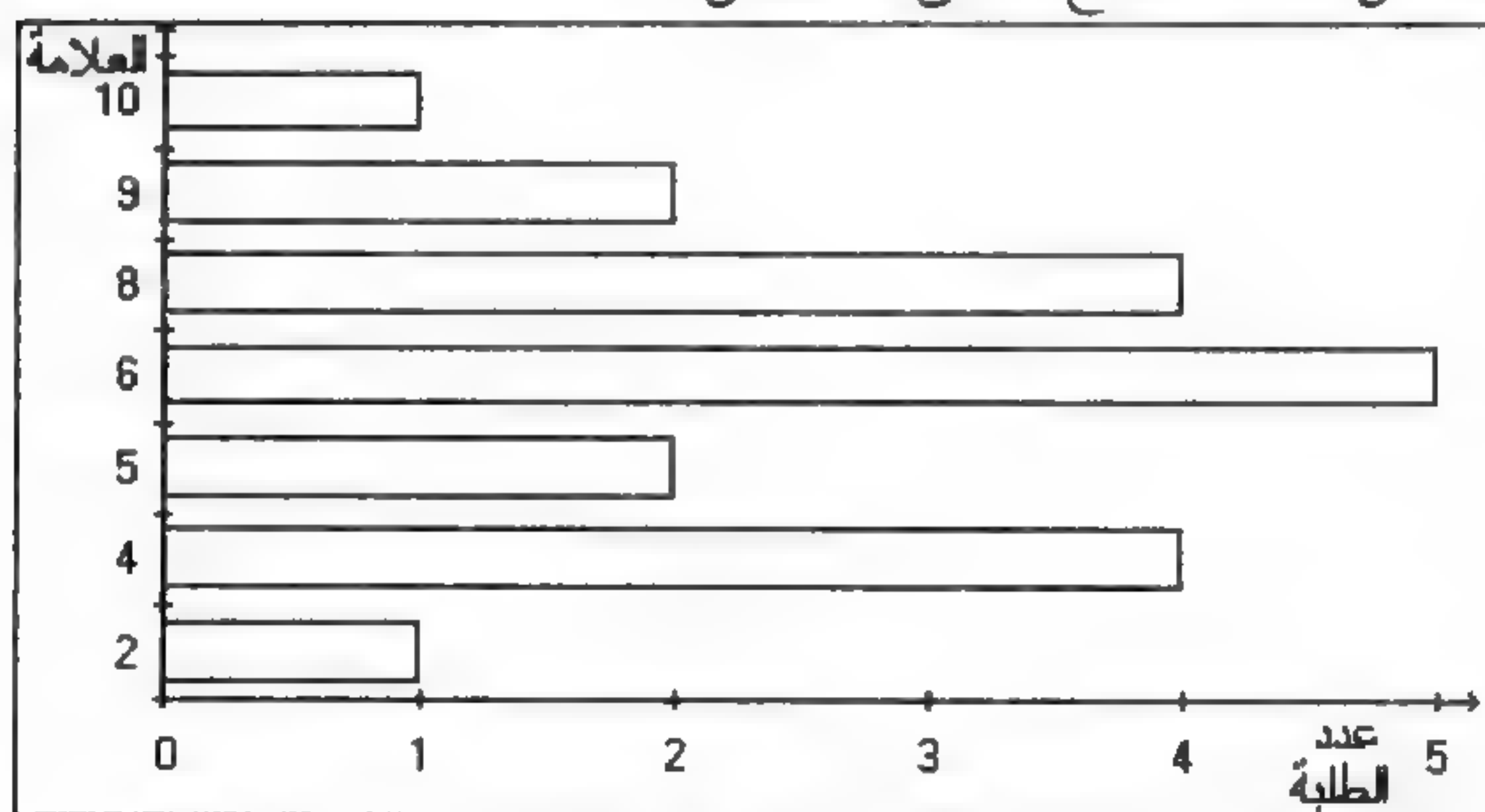
شكل 1-3

ويمكن تقديم مثل هذه البيانات بأشكال مختلفة مع الحفاض على نفس المبدأ و من ذلك ما يلي:



شكل 2-3

كما يمكن قلبه ليصبح على الشكل :



شكل 3-3

فكل الأشكال السابقة ما هي إلا تجسيد لفكرة الأعمدة التكرارية، لكن بطرق مختلفة قليلا.

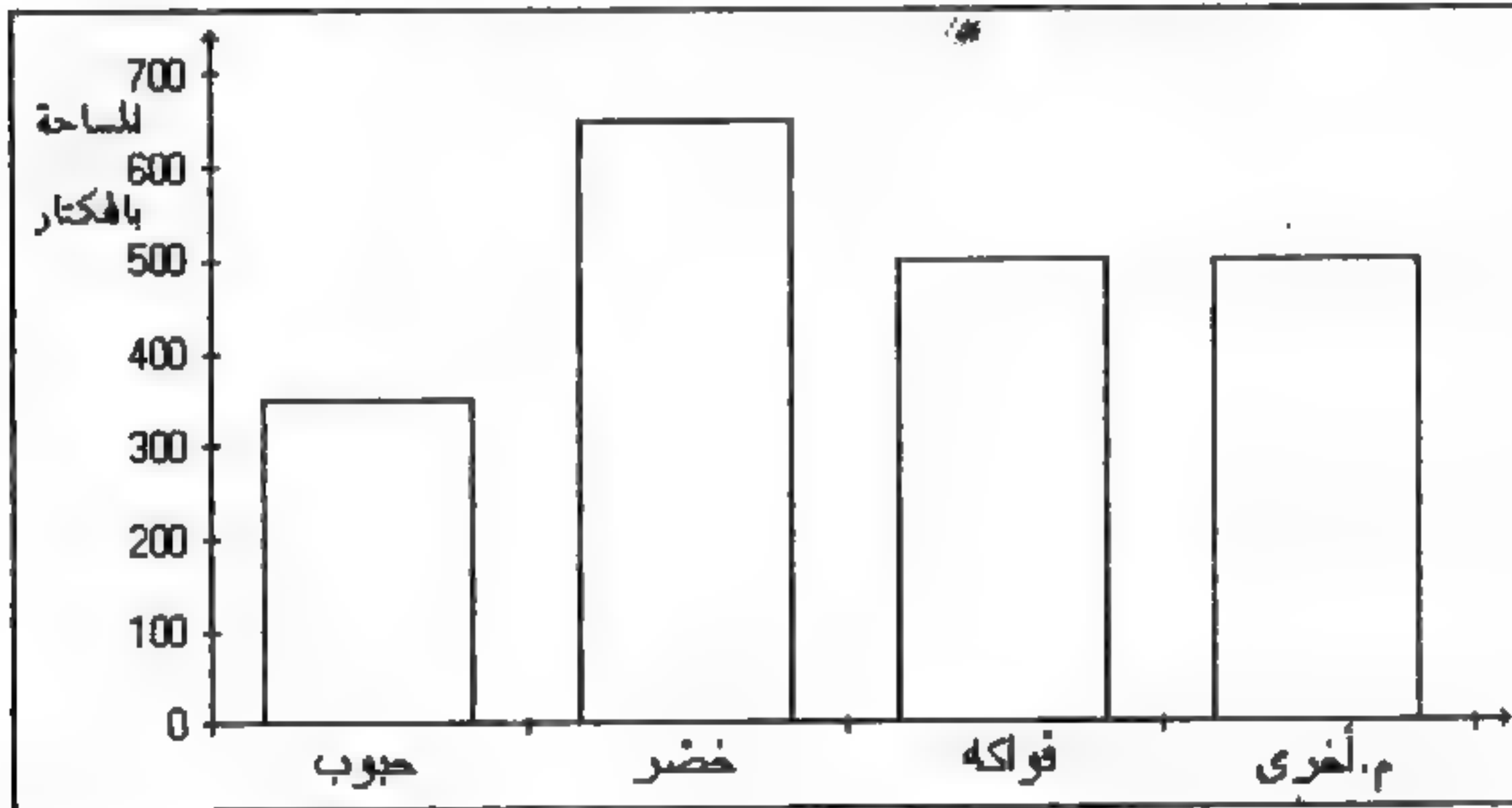
مثال 2-3: البيانات التالية تظهر توزيع مساحة مزرعة ما، حسب أنواع المزروعات خلال الموسم الفلاحي 2003/2004، والمطلوب تقديم هذه البيانات بطريقة الأعمدة التكرارية.

المساحة (هكتار)	المزروعات
350	الحبوب
650	الخضر
500	الفواكه
500	م. أخرى

جدول 2-3

كما سبقت الإشارة فإن أوصاف الظاهرة، وهي في مثالنا أنواع المزروعات، توضع على المحور الأفقي، أما المساحة فتوضع على المحور العمودي، ويكون الشكل المطلوب كما يلي :

توزيع مساحة مزرعة ما حسب أنواع المزروعات خلال الموسم 2004/2003



شكل 3-4

ويمكن التفنن في تقديم مثل هذه البيانات، كما في الشكلين 3-3 و 2-3، أو برسم أنواع المزروعات في الأعمدة المناسبة في الشكل 3-4، كأن نرسم في عمود الفواكه حبة برتقال، أو تفاح... وفي عمود الحبوب سنبلة قمح أو شعير... ونستغني بذلك عن الكتابة تحت الأعمدة، أي أن نستعمل الأسلوب الملفت للإنتباه والذي يؤدي إلى فهم وتحليل محتوى الشكل بأسرع ما يمكن.

2- **المدرج التكراري** : يستخدم في غالب الأحيان لتقديم البيانات الإحصائية التي مدى فئتها أكبر من الصفرة إذ توضع الفئات الفعلية

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

على المحور الأفقي والتكرارات على المحور العمودي، وحيث أن الفئات التكرارية قد تكون متساوية الأطوال وقد تكون غير متساوية الأطوال، لذلك فهناك نوعين من المدرجات التكرارية:

• **مدرج البيانات التكرارية المتساوية الأطوال :** يتم رسمه على معلم متعامد حيث يوضع على المحور الأفقي الفئات، وعلى المحور العمودي التكرارات، مع تقسيم المحورين بحيث يكفي المحور الأفقي لتمثيل كل الفئات والمحور العمودي لتمثيل كل التكرارات (يتم ذلك بتقسيم أكبر قيمة من مجموعة القيم على طول المحور المعد للرسم، حيث نحصل على القيم التي تمثلها كل سلمية على المحور)، و نقوم بعد ذلك برسم مستطيلات، ضلعها الأفقي يساوي طول الفئة وضلعها العمودي يساوي قيمة تكرار تلك الفئة، وذلك بالنسبة لجميع الفئات، والمثال التالي يوضح ذلك:

مثال 3-3 : البيانات التالية تظهر توزيع الموظفين في القطاع العمومي في مدينة ما حسب فئات الأعمار، بآلاف الأشخاص.

1	فئات الأعمار (سنة)	العدد
1	25-20	9
2	30-25	10
3	35-30	15
4	40-35	15
5	45-40	10
6	50-45	9
7	55-50	9
8	60-55	6
9	65-60	4

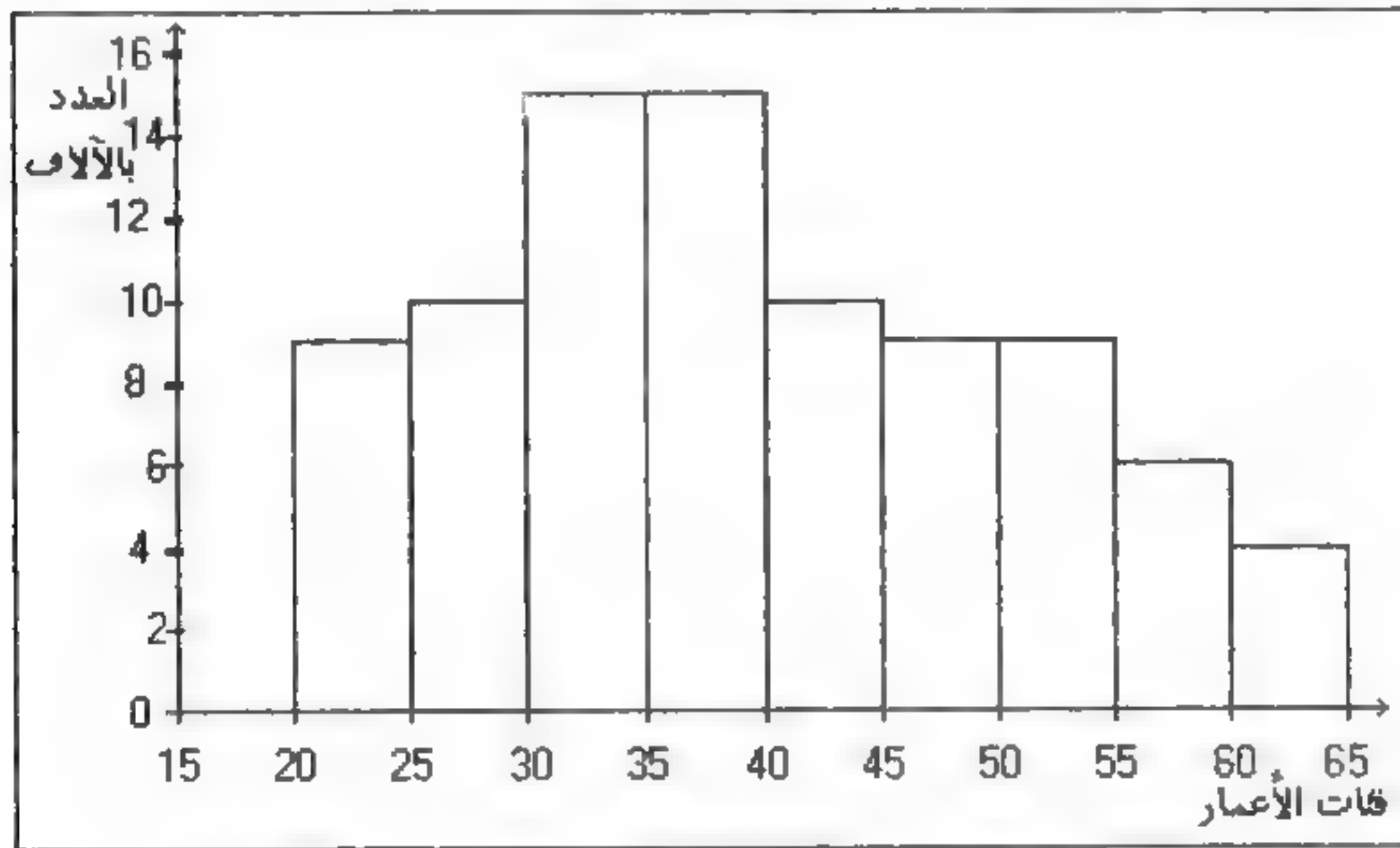
جدول 3-3

المطلوب : تقديم هذه البيانات في شكل مدرج تكراري.

الإجابة : أكبر قيمة من قيم التكرارات هي 15 بينما طول المحور العمودي المعد للرسم هو 4 سم، لذلك فإن السلمية الواحدة من هذا المحور تساوي : $4 \setminus 15 = 4$ تقريباً.

بينما أكبر قيمة للفئات هي 65 و طول المحور المعد للرسم هو : 13 سم، لذلك فإن طول السلمية الواحدة يساوي الى : $13 \setminus 65 = 5$.

وبتطبيق فكرة الرسم أعلاه نحصل على الشكل التالي:
توزيع الموظفين في القطاع العمومي في مدينة ما حسب فئات الأعمار



شكل 3-5

من الواضح أن الشكل أعلاه يشبه المدرج، لذلك سمي بالمدرج التكراري. يلاحظ من الشكل 3-5 أن مساحة المستطيلات تتناسب مع تكراراتها، فإذا كانت قاعدة المستطيلات هي L (طول الفئة) وإرتفاعها هو y_i (تكرار الفئة) فإن:

$$\frac{\text{مساحة المستطيل 1}}{f_1} = \frac{\text{مساحة المستطيل 2}}{f_2} = \frac{L_1 y_1}{f_1} = \frac{L_2 y_2}{f_2} \quad 1-3$$

و هو المبدأ الذي تقوم عليه فكرة المدرجات التكرارية، و بما أن أطوال الفئات متساو، فيبقى التناسب فقط بين ارتفاع المستطيلات والتكرارات أي:

$$\frac{y_1}{f_1} = \frac{y_2}{f_2} = \dots$$

* مدرج البيانات التكرارية غير المتساوية

الأسوال: في هذه الحالة ولجعل المدرج متزنا من حيث الشكل، فإنه يتم أولا إجراء تعديل على التكرارات، وذلك بإيجاد ما يسمى بالتكرار المعدل و ذلك بإحدى الطريقتين:

الطريقة الأولى: طريقة النسبة الى طول الفئة : (أكثر شيوعا).
كما هو واضح من اسم هذه الطريقة فإنه يتم إيجاد ما يسمى بالتكرار المعدل وذلك بقسمة تكرار كل فئة على طول الفئة المقابلة أي :

$$fcl_i = \frac{f_i}{L_i} \quad 2-3$$

حيث: fcl_i التكرار المعدل بالطريقة الأولى للفئة i مع $i = 1, 2, 3, \dots, N$
, و يتم وضع التكرار المعدل على المحاور العمودي أما الفئات فيتم وضعها على المحور الأفقي، ويتم الرسم بعد ذلك بصفة عادية.
القاعدة هي أنه عند رسم المصّلع التكراري يجب الحفاظ على التناسب بين التكرارات المعدلة وإرتفاع المستطيلات، وهذا هو المبدأ الذي تقوم عليه هذه الطريقة.

الطريقة الثانية: طريقة النسبة الى مضاعف أدنى طول فئة: كما هو واضح أيضا من إسم هذه الطريقة فإنه يتم إيجاد التكرار المعدل وذلك بقسمة تكرار كل فئة على مضاعف أدنى طول فئة في التوزيع، فإذا كانت لدينا بيانات تكرارية مستمرة، أطوالها على التوالي:

$$c_1.L, c_2.L, c_3.L, \dots, c_n.L$$

حيث L أصغر طول من أطوال الفئات ، c_i : معامل (حاصل قسمة طول الفئة i على L).

وتكراراتها على التوالي هي : $f_1, f_2, f_3, \dots, f_n$

فان التكرار المعدل لكل فئة i يكتب بالصيغة التالية:

$$fc2_i = \frac{f_i}{c_i} \quad 3-3$$

حيث: $fc2_i$ التكرار المعدل بالطريقة الثانية للفئة i مع $i = 1, 2, 3, \dots, N$
 ولرسم المدرج يتم وضع الفئات على المحور الأفقي والتكرارات المعدلة على المحور العمودي، و يجري الرسم بصفة عادية.
 تقوم هذه الطريقة أيضا على مبدأ التناسب بين مساحة المستطيلات والتكرارات، أي أنه عند رسم المدرج التكراري بهذه الطريقة ينبغي مراعاة التناسب بين التكرارات المعدلة وارتفاعات المستطيلات.

مثال 3-4: ارسم المدرج التكراري للبيانات التالية :

6	5	4	3	2	1	i
65-55	55-40	40-35	35-25	25-15	15-10	الفئة
10	27	15	24	16	6	f_i

جدول 3-4

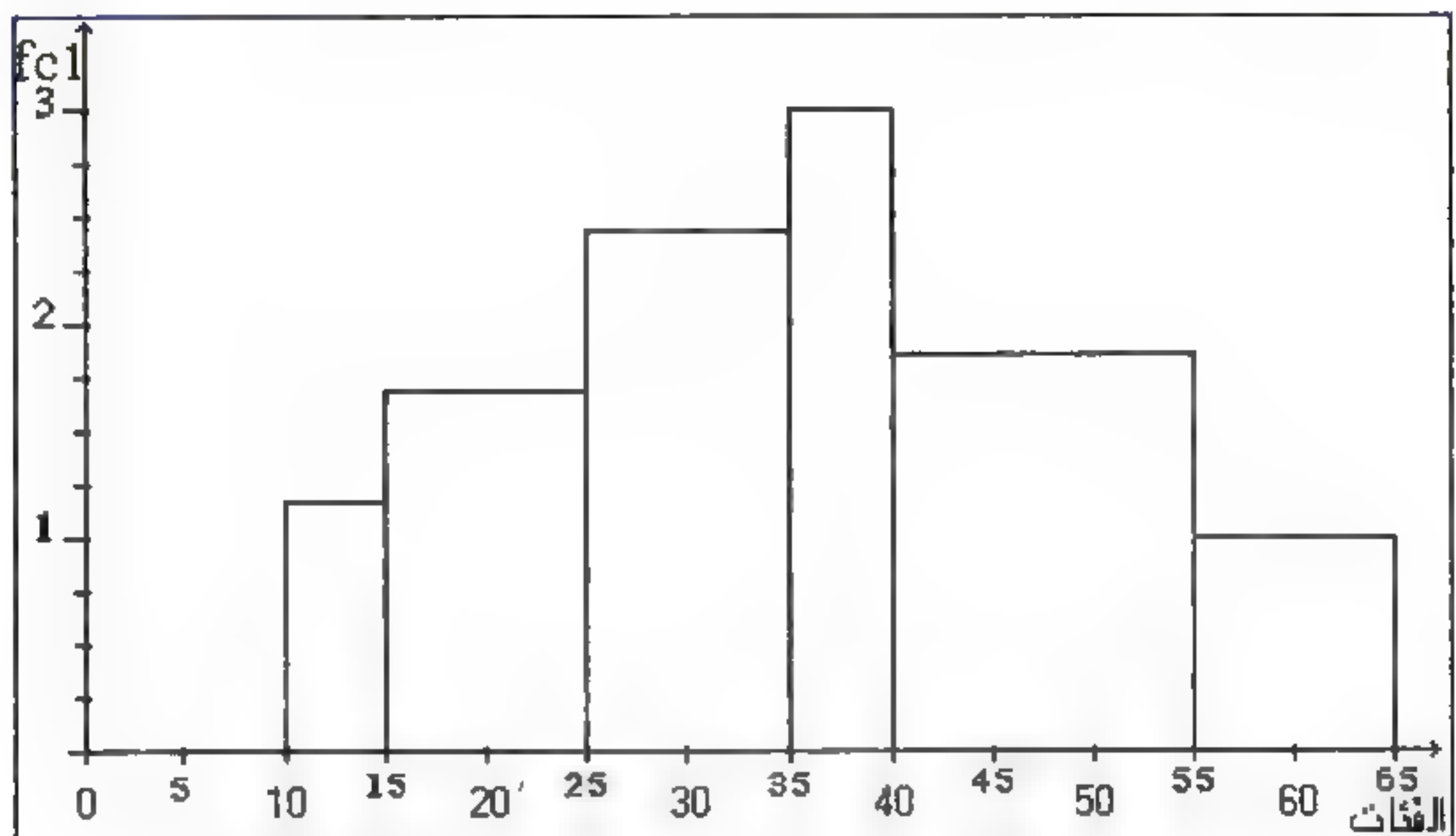
الإجابة: من الجدول أعلاه يلاحظ أن أطوال الفئات غير متساوية، لذلك يتم استخدام التكرار المعدل المعروف بإحدى الطريقتين المشار إليهما أعلاه، والمحسوب في الجدول 3-5 أدناه، اذ يحتوي على التكرار المعدل بالطريقة الأولى في العمود 5 وعلى التكرار المعدل بالطريقة الثانية في العمود 8.

$fc2_i$	c_i	$c_i \cdot L$	$fc1_i$	L_i	f_i	الفئة	i
6	1	1×5	1.2	5	6	15-10	1
8	2	2×5	1.6	10	16	25-15	2
12	2	2×5	2.4	10	24	35-25	3
15	1	1×5	3	5	15	40-35	4
9	3	3×5	1.8	15	27	55-40	5
5	2	2×5	1	10	10	65-55	6

جدول 3-5

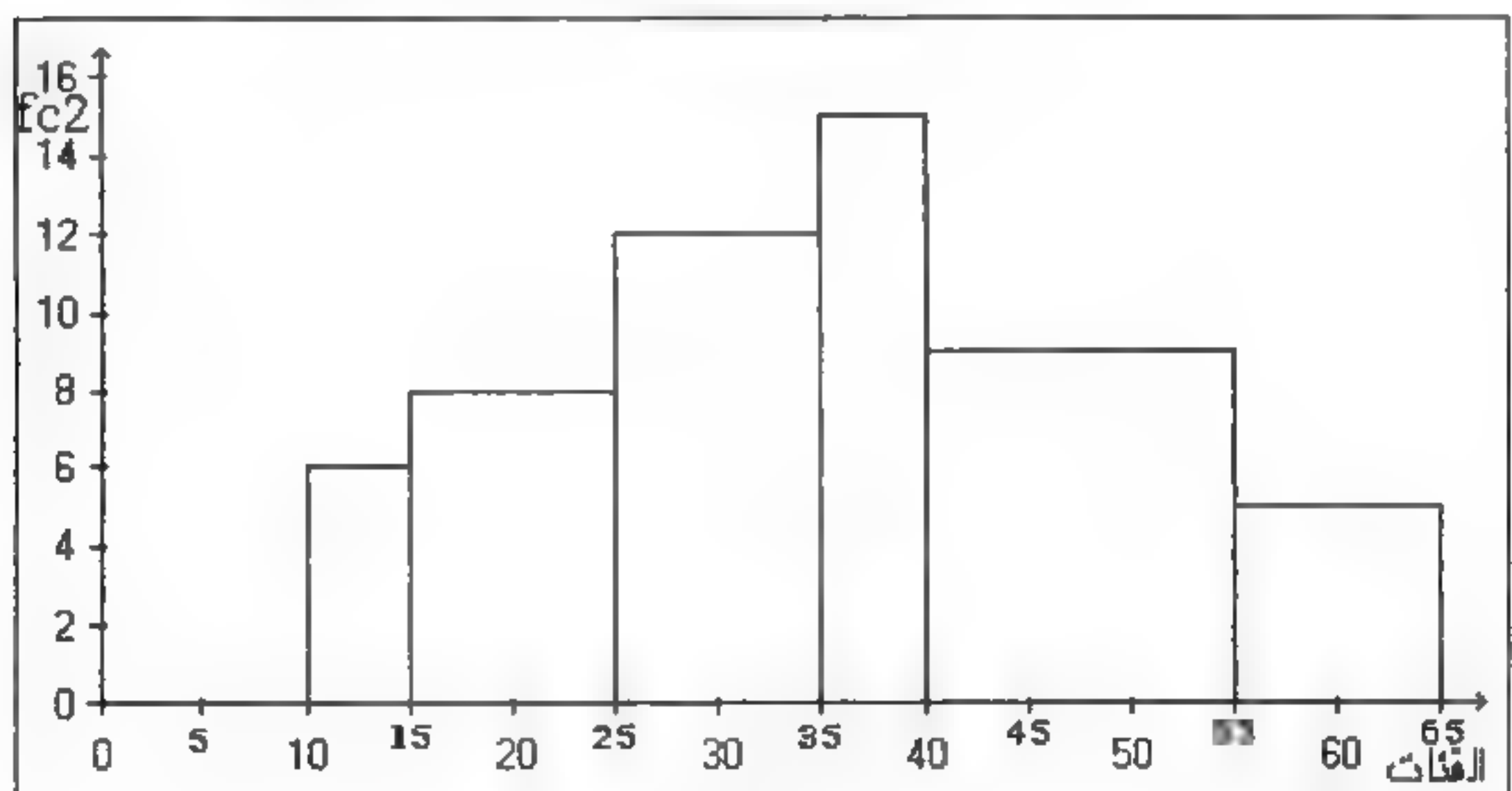
ويكون المدرج التكراري المطلوب بالطريقة الأولى والثانية في الشكلين 3-6 و 3-7 على التوالي هو :

مدرج تكراري معدل لبيانات المثال 3-4 بالطريقة الأولى.



شكل 3-6

مدرج تكراري معدل لبيانات المثال 3-4 بالطريقة الثانية.



شكل 3-7

3- المضلع التكراري : ويغلب إستخدامه أيضا في حالة البيانات التي طول فئاتها أكبر من الصفر، ويتم ذلك حسب الخطوات التالية :

* نرسم معلم متعامد، نضع على محوره الأفقي مراكز الفئات، وعلى محوره العمودي التكرارات، حيث : (مركز الفئة = الحد الأعلى للفئة + الحد الأدنى لها مقسوما على 2) أي :

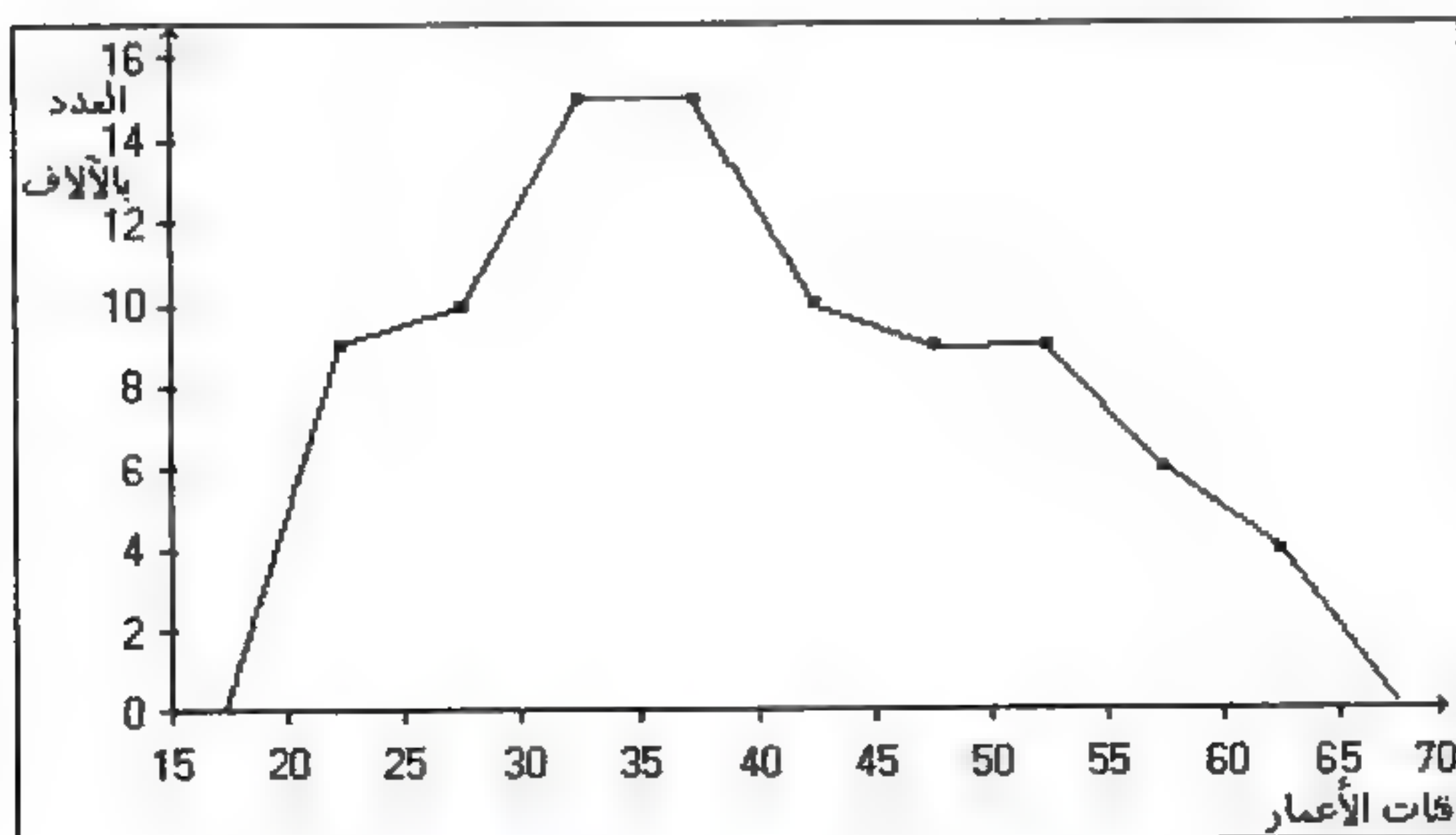
$$c_i = \frac{T_{i+1} + T_i}{2} \quad 4-3$$

حيث : c_i مركز الفئة i ، T_{i+1} : الحد الأعلى للفئة i ، T_i : الحد الأدنى للفئة i .
* نعين النقاط التي إحداثياتها مراكز الفئات والتكرارات المقابلة لها، ثم نصل بينها فنحصل على خط منكسر.
* لكي نحصل على مضلع، يتم إحداث مركز فئة تصوري سابق لأول فئة، وآخر لاحق لآخر فئة، تكراراتهما معدومة، ثم نصل طرفي الخط المنكسر بهتين النقطتين.

مثال 3-5: ارسم مضلع تكراري لبيانات المثال 3-3.

الإجابة: لرسم المضلع التكراري المطلوب نتبع الخطوات المشار إليها أعلاه فنحصل على الشكل التالي :

توزيع الموظفين في القطاع العمومي في مدينة ما حسب فئات الأعمار

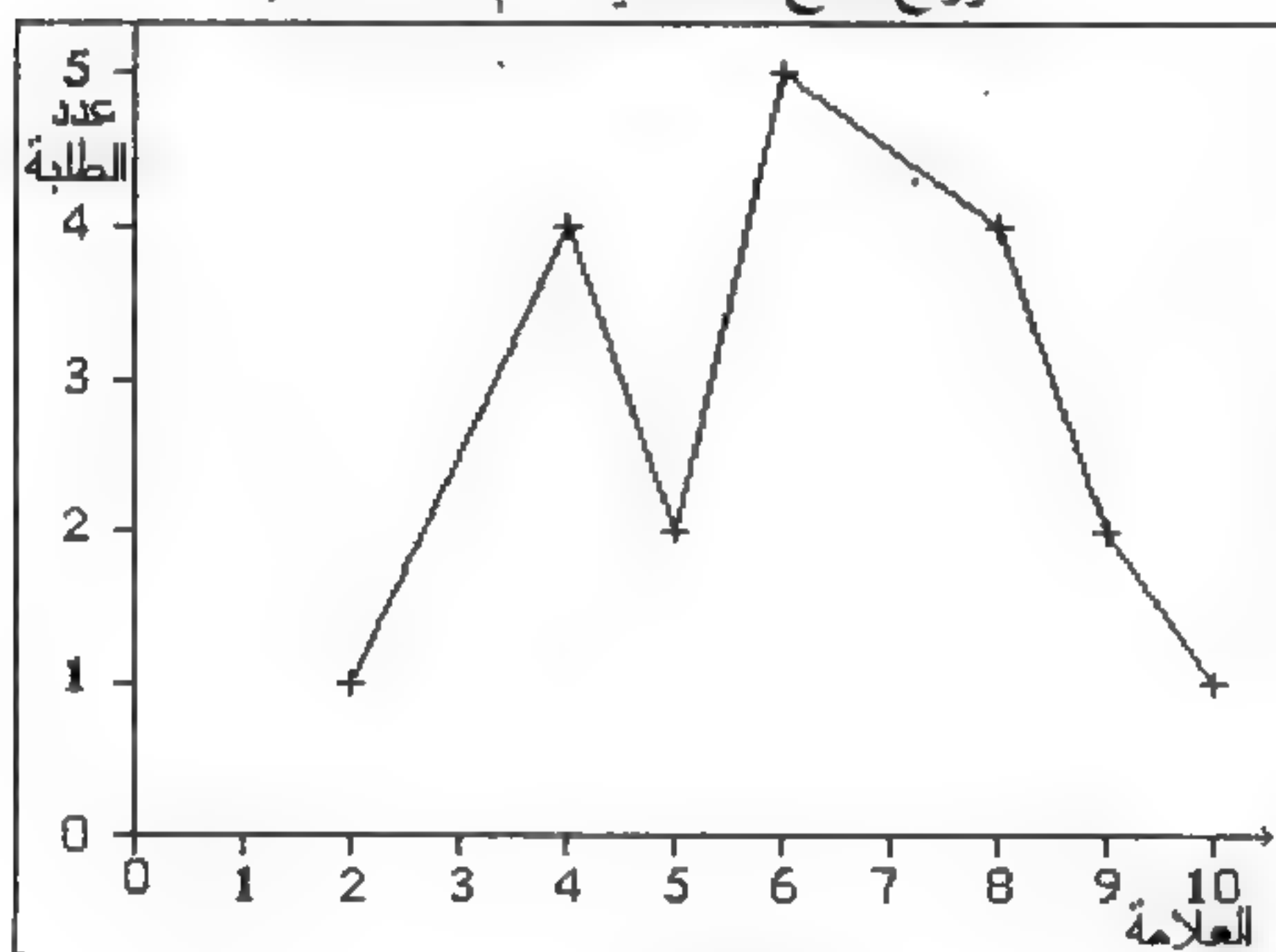


شكل 3-8

4- **المنحنى التكراري** : وهو المنحنى المحصل عليه بالوصل بين الفئات وتكراراتها في حالة البيانات التكرارية غير المستمرة، وبين النقاط المحددة لمراكز الفئات وتكراراتها في حالة البيانات التكرارية المستمرة.

مثال 3-6: إرسم المنحنى التكراري لبيانات المثال 3-1. البيانات المشار إليها بيانات تكرارية غير مستمرة، منحناها التكراري يكون حسب الشكل 3-9.

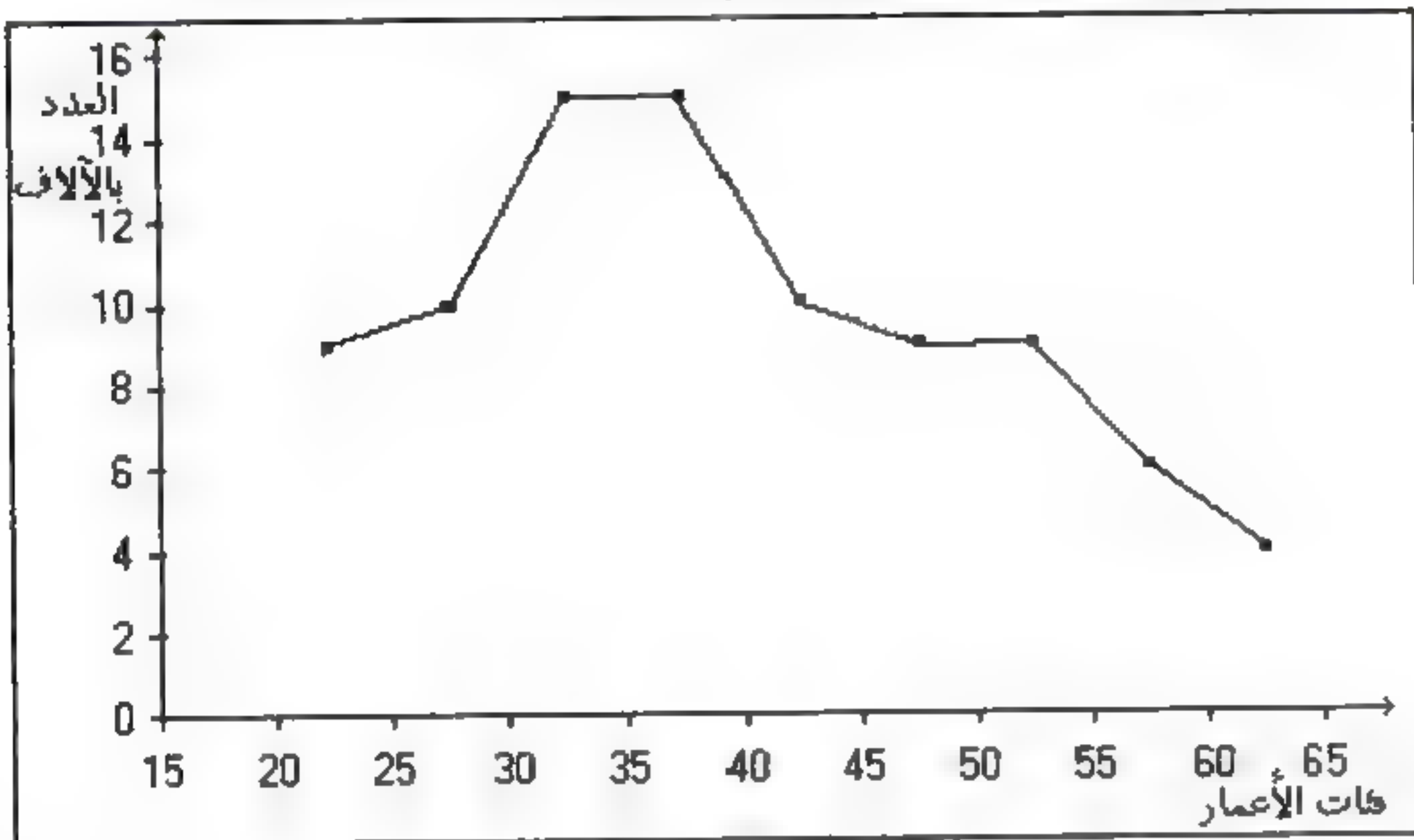
توزيع نتائج الطلبة في قسم به 19 طالب.



شكل 3-9

مثال 3-7: ارسم المنحنى التكراري لبيانات المثال 3-3. **الإجابة:** البيانات المشار إليها هي بيانات تكرارية مستمرة، منحناها التكراري يكون حسب الشكل 3-10، محوره الأفقي يمثل مراكز الفئات، و محوره العمودي يمثل التكرارات.

توزيع الموظفين في القطاع العمومي في مدينة ما حسب فئات الأعمار



شكل 3-10

وقد يتم إستبدال الخطوط المستقيمة الواصلة بين النقاط بخطوط ممهدة.

5- **المدرج التكراري المتجمع** : ويمكن أن يكون إما صاعداً أو نازلاً (أنظر التوزيع التكراري المتجمع الصاعد و التوزيع التكراري المتجمع النازل في الفصل السابق).

* **المدرج التكراري المتجمع الصاعد** : ويتم تقديمه على معلم متعامد، بحيث توضع الفئات على المحاور الأفقي والتكرارات المتجمعة الصاعدة على المحور العمودي، ويتم رسم المدرج بشكل مشابه للمدرج التكراري كما ورد آنفاً.

* **المدرج التكراري المتجمع النازل** : لتقديم البيانات بهذا الشكل يتم إيجاد التكرار المتجمع النازل، ثم رسم معلم متعامد، بحيث نضع على محوره الأفقي الفئات وعلى محوره العمودي التكرارات المتجمعة النازلة، ويتم رسم المدرج أيضاً بشكل مشابه للمدرج التكراري.

مثال 3-8: قدم البيانات التالية مرة في شكل مدرج تكراري متجمع صاعد، وأخرى في شكل مدرج تكراري متجمع نازل.

f_i	الفترة	i
5	3-2	1
8	4-3	2
12	5-4	3
6	6-5	4
4	7-6	5
35		مج

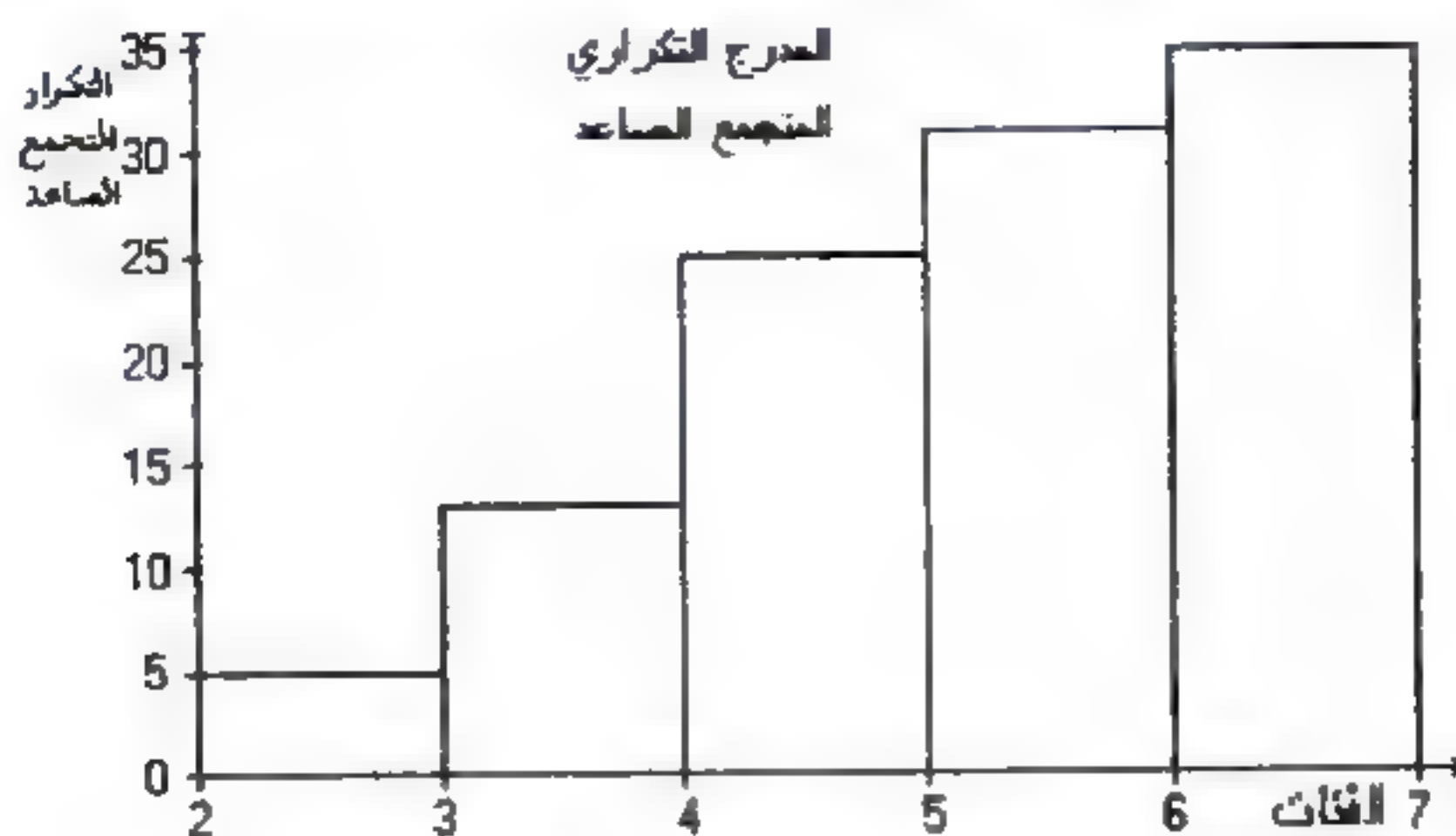
جدول 3-6

* المدرج التكراري المتجمع الصاعد : لرسمه نوجد التكرار المتجمع الصاعد كما هو واضح أدناه :

f_i	الفترة	i	الحد الأعلى
5	3-2	1	أقل من 3
13	4-3	2	أقل من 4
25	5-4	3	أقل من 5
31	6-5	4	أقل من 6
35	7-6	5	أقل من 7
35		مج	

جدول 3-7

ويكون الرسم المطلوب كمالى:



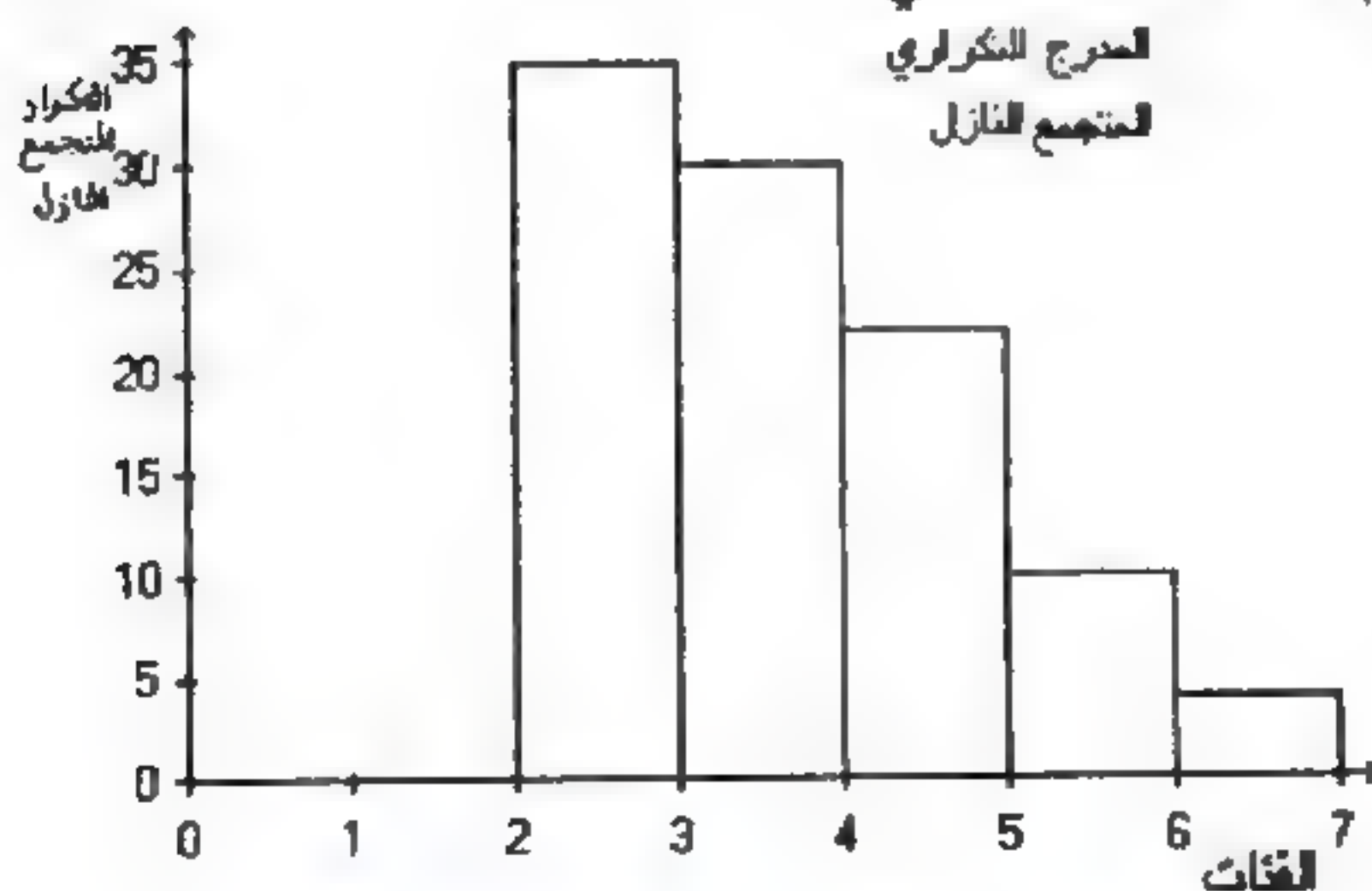
شكل 3-11

* المخطط التكراري المتجمع النازل : لرسمه نوجد التكرار المتجمع النازل كما يوضحه الجدول التالي :

i	الفئة	f_i	الحد الأدنى	
1	3-2	5	2 فأكثر	35
2	4-3	8	3 فأكثر	30
3	5-4	12	4 فأكثر	22
4	6-5	6	5 فأكثر	10
5	7-6	4	6 فأكثر	4
مج		35		

جدول 3-8

ويكون الرسم المطلوب كما يلي :



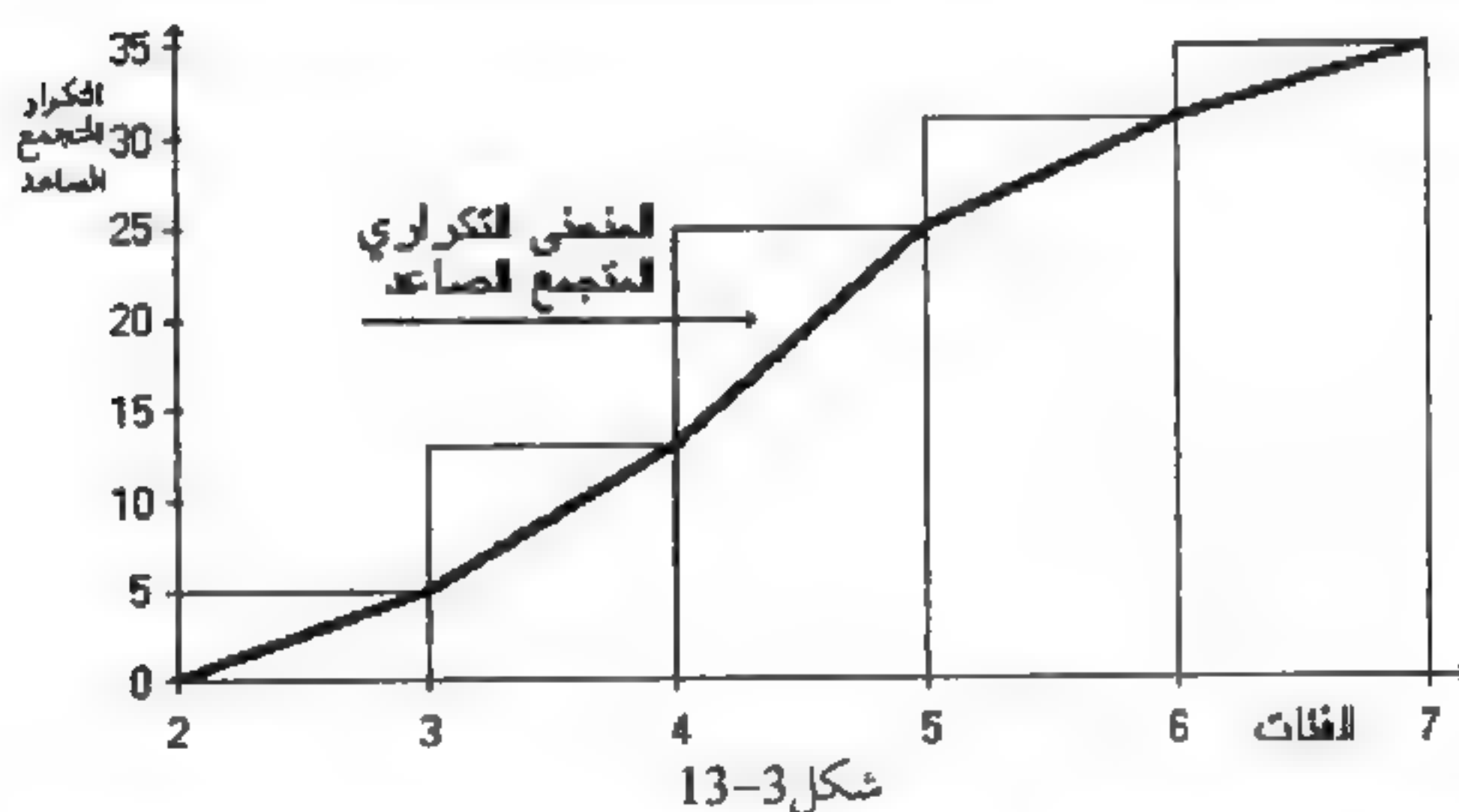
شكل 3-12

6- المنحنى التكراري المتجمع : وقد يكون أيضا إما متجمعا صاعدا، أو متجمعا نازلا :

* **المنحنى التكراري المتجمع الصاعد :** ويتم رسمه انطلاقا من المدرج التكراري المتجمع الصاعد، وذلك بالوصل بين النقاط التي تمثل الحدود العليا للفئات والتكرارات المقابلة لها في المدرج التكراري المتجمع الصاعد، كما هو واضح في المثال أدناه.

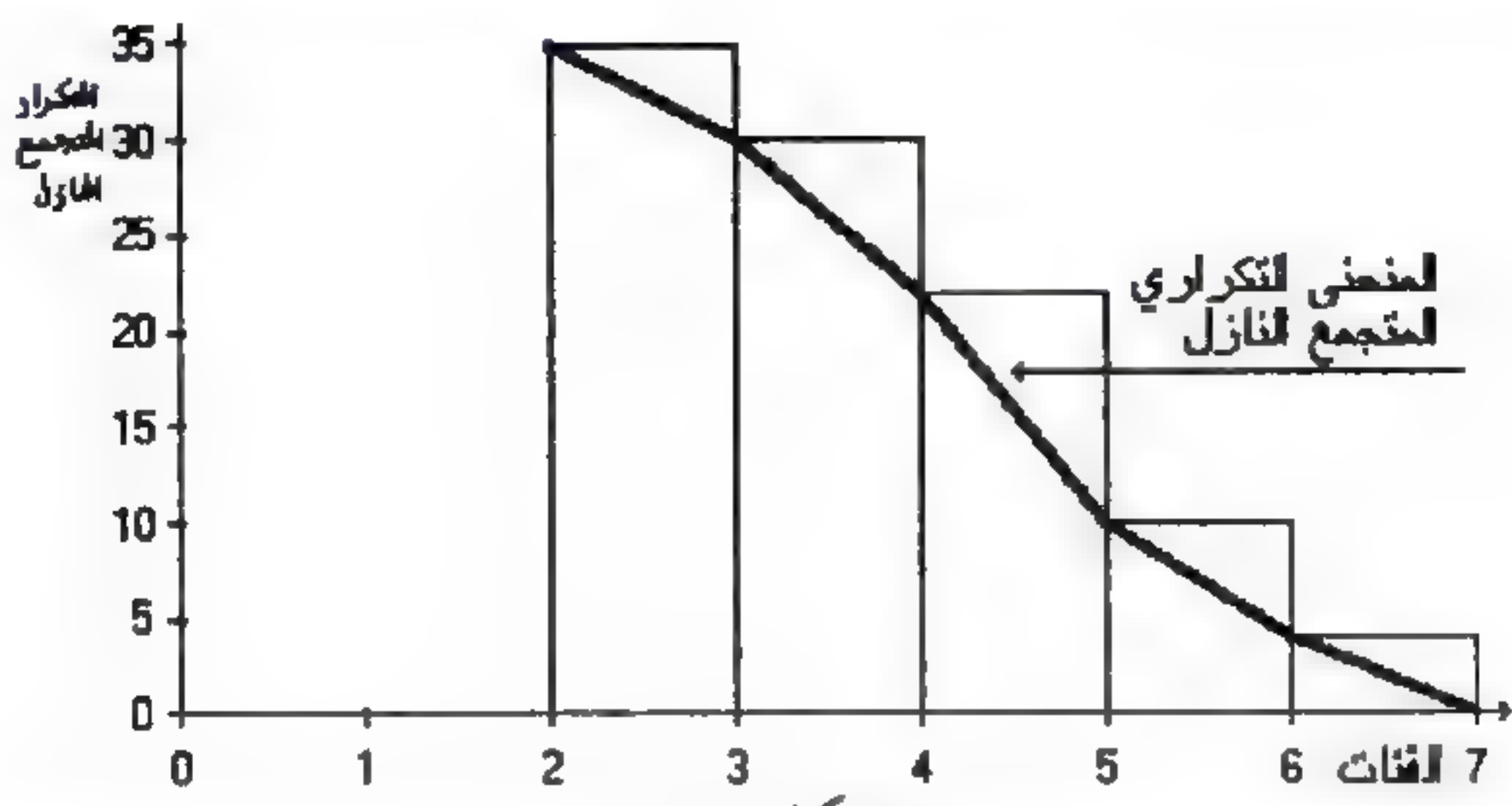
مثال 3-9: ارسم المنحنى التكراري المتجمع الصاعد لبيانات المثال 3-8.

بتطبيق مبدأ الرسم المشار إليه أعلاه نحصل على الشكل المطلوب :



* **المنحنى التكراري المتجمع النازل :** ويتم رسمه انطلاقا من المدرج التكراري المتجمع النازل، حيث يتم الوصل بين النقاط التي تمثل الحدود الدنيا للفئات والتكرارات المتجمعة المقابلة لها، كما هو واضح في المثال التالي :

مثال 3-10: ارسم المنحنى التكراري المتجمع النازل لبيانات المثال 3-8. بتطبيق مبدأ الرسم المشار إليه أعلاه نحصل على الشكل المطلوب.



شكل 3-14

ملاحظة : يمكن تقديم كل الأشكال التكرارية حسب الترتيب الذي سقناه، عن طريق الرسومات التكرارية النسبية (المئوية)، وذلك باستبدال التكرارات المطلقة على المحاور العمودية بالتكرارات النسبية (المئوية).

ثالثا: المنحنيات الزمنية : وهي المنحنيات التي تظهر تطور ظاهرة ما عبر الزمن، سواء كان مقاسا بالأيام أو بالسنوات أو بأجزائهما أو أضعافهما، في هذه الحالة يتم رسم معلم متعامد، يوضع على محوره الأفقي الزمن وعلى محوره العمودي قيمة الظاهرة، ثم تحدد نقاط التوفيقات المختلفة بين الزمن وقيمة الظاهرة، ليوصل بينها، فنحصل بذلك على المنحنى المطلوب.

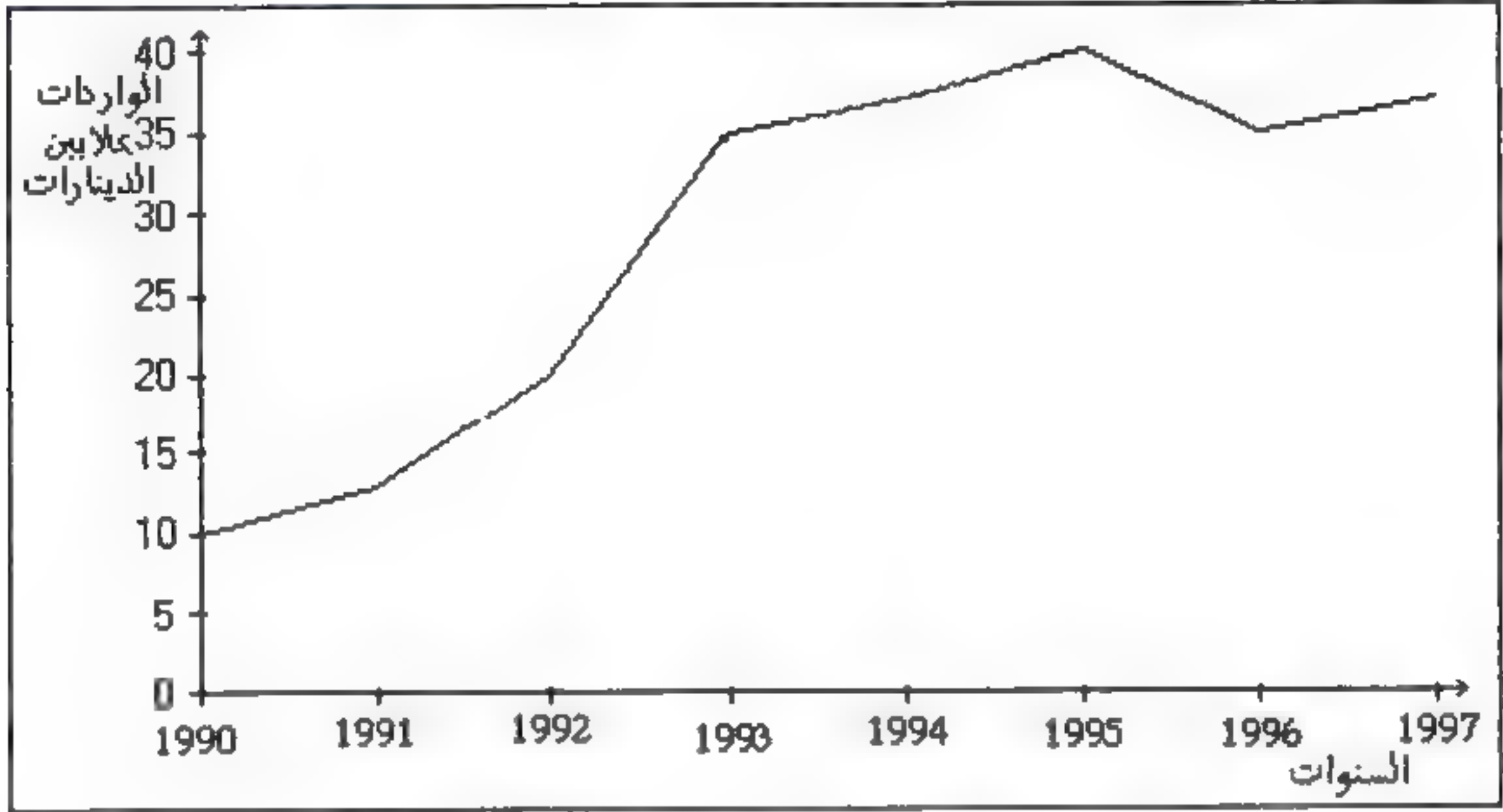
مثال 3-11: البيانات التالية تظهر تطور الواردات السلعية خلال الفترة: 1990 - 1997، بملايير الدينارات. المطلوب تقديمها في شكل بياني مناسب.

السنة	1990	1991	1992	1993	1994	1995	1996	1997
الواردات	10	13	20	35	37	40	35	37

جدول 3-9

يتم تقديم البيانات المشار إليها كما يلي :

تطور الواردات السلعية خلال الفترة : 1997-90



شكل 3-15

رابعاً: الشكل الدائري: يستخدم الشكل الدائري في غالب الأحيان، لتقديم بيانات ظاهرة ما، تتركب من مكونات جزئية خلال ظرف زماني أو مكاني محددين، كتوزيع كميات التساقط حسب الفصول في سنة ما، أو توزيع مساحة مزرعة ما حسب أنواع المزروعات... الخ، و تقدم البيانات عن طريق الشكل الدائري، يتم عن طريق النسبة الزاوية، حيث أن القيمة الكلية للظاهرة، تقابل الزاوية الكلية للدائرة أي 360° ، بينما القيم الجزئية للظاهرة تقابل أجزاء معينة من الزاوية الكلية للدائرة، يتم حسابها باستخدام القاعدة الثلاثية وذلك كما يلي :

360° ————— القيمة الكلية للظاهرة

x° ————— القيمة الجزئية للظاهرة

ومنه تكون الزاوية التي تقابل القيمة الجزئية للظاهرة كما يلي :

$$x^\circ = \frac{\text{القيمة الجزئية من الظاهرة}}{\text{القيمة الكلية للظاهرة}} \times 360$$

مثال 3-12: الجدول التالي يظهر المساحة المزروعة، حسب أنواع المزروعات في مستعمرة فلاحية مساحتها 1000 هكتار، خلال الموسم الفلاحي 1994، والمطلوب تقديمها عن طريق الشكل الدائري.

المزروعات	المساحة (هكتار)
حبوب	350
محضر	150
فواكه	200
م. أخرى	300

جدول 3-10

لرسم الشكل الدائري، نقوم بحساب الزوايا التي تقابل كل قيمة جزئية، باستخدام القاعدة 3-15 أعلاه حيث نجد:

المزروعات	المساحة	الزاوية المقابلة بالدرجات
حبوب	350	126
محضر	150	54
فواكه	200	72
م. أخرى	300	108
المجموع	1000	360

و بالتالي الشكل المطلوب هو:

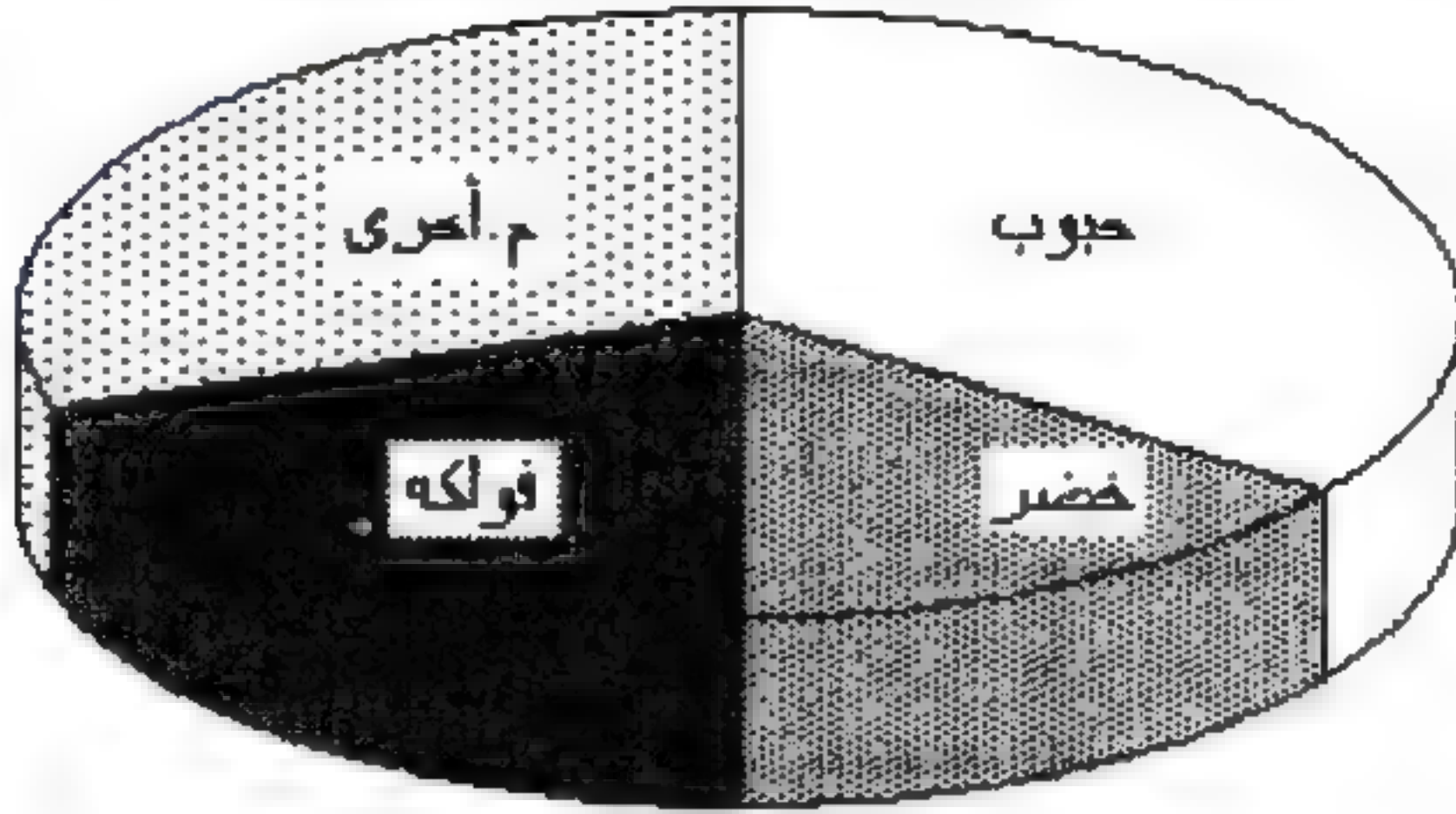
توزيع مساحة مزرعة ما حسب أنواع المزروعات.



شكل 3-16

ويمكن وضع الدائرة أيضا بشكل أفقي (أسطواني) على النحو:

توزيع مساحة مزرعة ما حسب أنواع المزروعات.



شكل 3-17

كما يمكن تقديم البيانات من هذا النوع أيضا عن طريق الشكل النصف دائري (على أساس 180°) أو الربع دائري (على أساس 90°).

خامسا: الشكل المستطيل: تقدم من خلاله البيانات المشابهة لتلك التي تقدم عن طريق الشكل الدائري، وفيه يجب أن تقابل القيمة الكلية للظاهرة المساحة الكلية للمستطيل المعد للرسم، و القيمة الجزئية للظاهرة، جزء من مساحة ذلك المستطيل، ويتم إيجادها عن طريق القاعدة الثلاثية كما يلي:

المساحة الكلية للمستطيل \rightarrow القيمة الكلية للظاهرة

$x \rightarrow$ القيمة الجزئية للظاهرة

ومنه نجد:

$$x = \frac{\text{القيمة الجزئية من الظاهرة}}{\text{القيمة الكلية للظاهرة}} \times \text{المساحة الكلية للمستطيل}$$

3-16

حيث: x : جزء من مساحة المستطيل.

لتسهيل الرسم ينصح أن يكون عرض المستطيل وحدة قياس واحدة، ليعتمد الرسم بذلك على طول المستطيل فقط.

مثال 3-13: قدم بيانات المثال 3-12 عن طريق الشكل المستطيل.

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

إذا فرضنا أن المساحة الكلية للمستطيل الذي نريد أن نقدم البيانات من خلاله هي: 10 سم²، فإنه يفضل أن يكون عرضه 1 سم وطوله 10 سم، وتجرى الحسابات بصفة مشابهة لما يلي :

1000 هكتار تقابل 10 سم²

350 هكتار (حبوب) تقابل x سم²

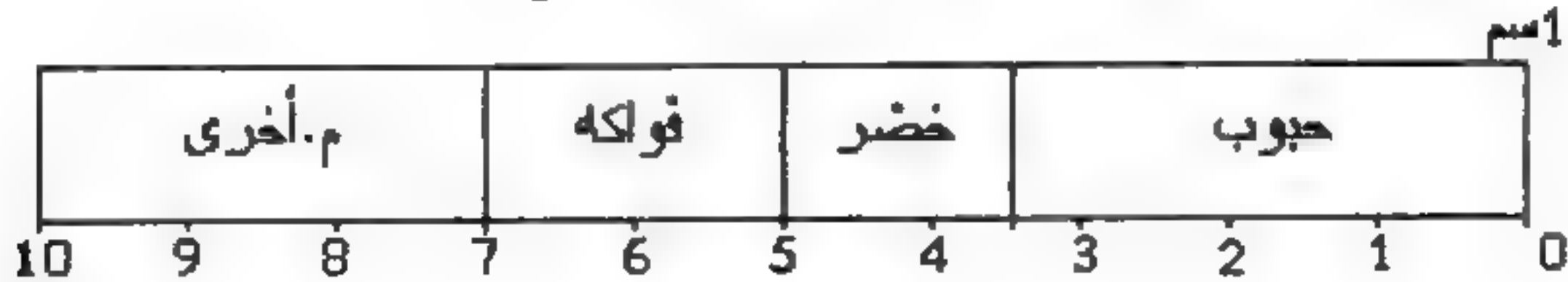
ومنه نجد أن المساحة المقابلة 350 هكتار من الحبوب، في مساحة المستطيل هي:

$$x = \frac{350}{1000} \times 10 = 3.5 \text{ سم}^2$$

وبالمثل نجد المساحة التي تقابل بقية أنواع المزروعات ضمن مساحة المستطيل، بتطبيق المعادلة 3-15 و هي:

المزروعات	المساحة (هكتار)	المساحة المقابلة على المستطيل سم ²
حبوب	350	3.5
خضار	150	1.5
فواكه	200	2.0
م. أخرى	300	3.0
المجموع	1000	10

و بالتالي يكون الشكل المطلوب كما يلي :



شكل 3-18

وتجدر الإشارة الى أنه يمكن تقديم البيانات التي تعطى على نفس منوال المثال 3-11 أعلاه، عن طريق الشكل المربع أيضا.

مادسا: الشكل القطبي: يستخدم هذا النوع من الأشكال عندما تكون البيانات الإحصائية خاصة بمواسم معينة أو أشهر خلال سنوات قليلة، ويتم ذلك حسب المثال التالي :

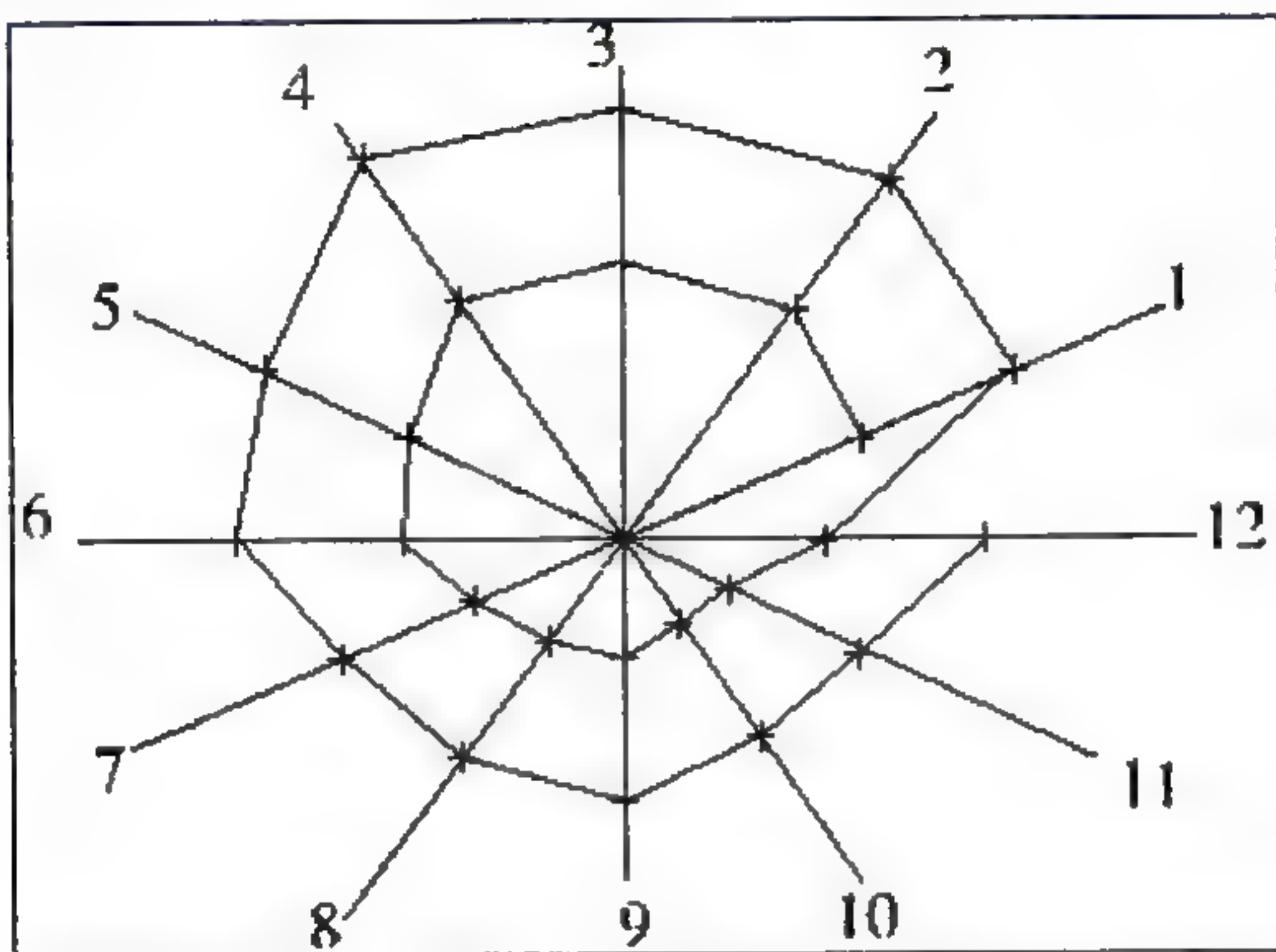
مثال 3-14: البيانات التالية تظهر تطور استهلاك مادة السكر في إحدى الولايات بآلاف الأطنان خلال أشهر سنتي 2001 و 2002.

سنة \ شهر	1	2	3	4	5	6	7	8	9	10	11	12
2001	20	22	23	24	22	20	19	18	18	16	16	18
2002	23	25	26	27	25	22	22	21	21	19	19	21

جدول 3-11

المطلوب : قدم هذه البيانات عن طريق الشكل القطبي.

الإجابة : نقسم مستوى الإحداثيين إلى 12 قسم متساوي، موصولة بنقطة الأصل بمستقيم كما هو واضح في الشكل 3-19، نعطي كل مستقيم رقم يدل على الشهر، ثم نختار وحدة قياس معينة، ونعين النقاط الموافقة لكل شهر على هذه المستقيمات ونصل بينها بخطوط في اتجاه عكس عقارب الساعة فنحصل بذلك على شكل يمثل البيانات كما هو أدناه:



شكل 3-19

مابعاً: الطريقة التصويرية: في هذه الطريقة يتم رسم عنصر الظاهرة بذاته، ليدل على عدد معين من عناصر تلك الظاهرة، فاذا كانت الظاهرة تتعلق بالأشجار المثمرة مثلاً، يتم رسم شجرة برتقال، أو شجرة تفاح، بحيث كل شجرة تمثل 1000 وحدة من الأشجار.

هذه هي أهم الأشكال البيانية التي تقدم عن طريقها البيانات الإحصائية، حسب طبيعتها، وهناك أشكال أخرى، غير أنها نادرة الاستعمال، إما لأن رسمها يتطلب حسابات معقدة نوعاً ما، أو لأن فهمها يكون صعباً. وكما سبقت الإشارة فإنه لتقديم أية بيانات عن طريق الرسومات، فإنه لابد من مراعاة قواعد الرسم، و الخواص الإحصائية المعروضة في بداية هذا البند.

و بتقديم البرامج المعلوماتية، فإن هناك الكثير من برامج الإعلام الآلي التي يمكن إستخدامها في تقديم البيانات الإحصائية، إذ يكفي الباحث بإدخال المعلومات، و يختار شكلاً من عشرات الأشكال المصممة، ليحصل على الشكل الذي يرغب فيه.

تمارين

- تمرين 1:** من التمرين الأول لسلسلة الفصل الثاني أجب:
- 1- ماهي الأشكال البيانية التي يمكن أن تقدم عن طريقها البيانات المحصل عليها بعد التبويب؟.
- 2- قدم البيانات المحصل عليها بكل الأشكال المناسبة.
- تمرين 2:** قدم النتائج المتوصل اليها من تبويب بيانات التمرين الثاني سلسلة تمارين الفصل الثاني بكل الأشكال البيانية الممكنة.
- تمرين 3:** ارسم المضلع التكراري و المضلع التكراري المسائوي والمنحنيين المتجمعين، الصاعد والنازل للبيانات المحصل عليها من التمرين الثالث سلسلة تمارين الفصل الثاني.
- تمرين 4:** البيانات التالية تظهر تطور سكان الجزائر خلال الفترة: 76-1986، بالملايين.

مصادر متعددة أساسها الديوان الوطني للإحصائيات

السنة	1976	1977	1978	1979	1980	1981
عدد السكان	16.3	17.1	17.7	18.1	18.7	19.2
السنة	1982	1983	1984	1985	1986	1987
عدد السكان	19.9	20.5	21.2	21.9	22.5	23.2
السنة	1988	1989	1990	1991	1992	1993
عدد السكان	23.8	24.4	25.0	25.6	26.3	26.9
السنة	1994	1995	1996	1997	1998	1999
عدد السكان	27.4	28.0	28.6	29.0	29.5	30.0
السنة	2000	2001	2002	2003	2004	2005
عدد السكان	30.4	30.8	31.2	31.6	32.1	

- المطلوب: 1-** أكمل الإحصائيات من مصادر الديوان الوطني للإحصائيات لغاية السنة التي نحن فيها.
- 2- قدم هذه البيانات بكل الأشكال البيانية الممكنة.

تمرين 5: البيانات التالية تظهر مساحات المحيطات بملايين الكيلومترات المربعة

المحيط	الهادي	الأطلسي	الهندي	م. الجنوبي	م. الشمالي
المساحة	183.4	106.7	73.8	19.7	12.4

المطلوب: قدم هذه البيانات بكل الأشكال البيانية الممكنة.

تمرين 6: البيانات التالية تظهر تطوّر عدد الناجحين في إمتحانات شهادة التعليم الأساسي حسب الجنس خلال الفترة 1998-2000 في الجزائر.

المصدر: الجزائر بالأرقام رقم 31. الديوان الوطني للإحصائيات. www.ons.dz

الجنس	جوان 1998	جوان 1999	جوان 2000
ذكور	105102	87767	110384
إناث	129077	106464	133221

المطلوب: 1- أوجد عدد المتحنيين في كل سنة

2- أوجد نسبة النجاح حسب الجنس. ماذا تستنتج.

3- قدم البيانات الأصلية و المحسوبة في السؤالين 1 و 2

بكل الأشكال البيانية الممكنة.

تمرين 7: البيانات التالية تظهر عدد المسجلين و الناجحين في شهادة البكالوريا لدورة جوان 2000 حسب الجنس في الجزائر.

المصدر: الجزائر بالأرقام رقم 31. الديوان الوطني للإحصائيات. www.ons.dz

	ذكور	إناث	المجموع
المسجلون	195147	250321	445486
الناجحون	49133	70192	119325

المطلوب: 1- أوجد نسبة النجاح من المجموع لكل فئة. ماذا تستنتج. حلل.

2- أوجد نسب النجاح حسب الجنس، ماذا تستنتج؟

3- قدم البيانات الأصلية و البيانات المحسوبة في السؤال 1 و 2

بكل الأشكال البيانية الممكنة

تمرين 8: من التمرين رقم 5 في سلسلة تمارين الفصل السابق:
1- إرسم الهرم السكاني للبيانات الخاصة بتوزيع السكان حسب فئات الأعمار لسنة 2000 ، في حالة $L=5$ ثم في حالة $L=10$.

2- أجب على نفس السؤال 1 بالنسبة للبيانات الخاصة بتوزيع السكان حسب فئات الأعمار لسنة 1993 في حالة $L=5$ ثم في حالة $L=10$.
3- قارن بين الهرمين. ماذا تستنتج.

تمرين 9: من بيانات التمرين 4 سلسلة تمارين الفصل الثاني قم بما يلي:

- 1- قدم مساحة الوطن العربي بالشكل البياني المناسب.
- 2- أوجد الكثافة السكانية لكل دولة و قدمها بالأشكال البيانية المناسبة.
- 3- قدم مساحة دول المغرب العربي بالشكل الدائري، النصف دائري، المستطيل، الأعمدة التكرارية.
- 4- نفس السؤال بالنسبة لدول الخليج العربي، و لدول الساحل الغربي للبحر الأحمر.

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

الفصل الرابع

مقاييس التزعة المركزية.

بعد جمع البيانات الاحصائية حول الظاهرة المدروسة، ينتقل الاحصائي الى دراستها وتحليلها وإستخلاص النتائج، ولأجل ذلك يكون من بين أولى إهتماماته بحث مدى تمركز القيم التي جمعها حول قيمة ما ضمن مجموعة البيانات الاحصائية، ومعنى آخر بحث مدى نزوع مختلف قيم الظاهرة حول قيمة مركزية منها، ويتم ذلك عن طريق أدوات تحليلية، سميت لأجل ذلك بمقاييس التزعة المركزية، وهي:

1-الوسط الحسابي. 4-الوسط التريعي.

2-الوسيط. 5-الوسط الهندسي.

3-النوال. 6-الوسط التوافقي.

يهدف هذا الفصل الى تبيان أهمية هذه المقاييس في التحليل الإحصائي، وكيفية حسابها، مع إبراز خصائصها والعلاقة فيما بينها، وقبل الشروع في ذلك نستخدم على ما يلي:

1-بيانات غير موزونة : هي البيانات غير المرتبة في شكل جداول تكرارية.

2-بيانات موزونة طول فئاتها معدومة : هي البيانات المرتبة في شكل جداول تكرارية غير مستمرة، أي طول فئاتها معدوم ($L = 0$)

3-بيانات موزونة مدى فئاتها أكبر من الصفر : هي البيانات المرتبة في شكل جداول تكرارية مستمرة، أي طول فئاتها أكبر من الصفر ($L > 0$).

قبل الشروع في تقديم مختلف مقاييس التزعة المركزية نشير الى أن أية قيمة مركزية تكون مثلى إذا توافرت فيها شروط يول Yule التالية:

- أن تؤخذ هذه القيمة دون تحيز من طرف الباحث.
- أن تحسب اعتمادا على جميع معطيات السلسلة الإحصائية.
- أن يكون لها معنى مادي.
- أن تكون سهلة الحساب.
- أن لا تتأثر كثيرا بتغيرات العينة أو المجتمع.
- أن تكون قابلة لإجراء الحسابات الجبرية عليها.

أولا: الوسط الحسابي : الوسط الحسابي لأية مجموعة من القيم هو معدلها بالتعبير العام، وتختلف طريقة حسابه حسب طبيعة البيانات كما هي موضحة أعلاه.

1- الوسط الحسابي للبيانات غير المبوبة :

تعريف 1-4: إذا كانت لدينا القيم $x_1, x_2, x_3, \dots, x_n$ ، فإن الوسط الحسابي لها يعطى بالمعادلة التالية:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{N} \quad 1-4$$

ويكتب اختصارا كمايلي :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} \quad 2-4$$

حيث : $i=1, 2, 3, \dots, N$ (هو عدد القيم)

ويعني هذا أن الوسط الحسابي للبيانات غير المبوبة يساوي الى مجموع البيانات مقسوما على عددها.

مثال 1-4: حصل طالب على النتائج التالية في 10 مقاييس :

12 13 14 12 10 12 13 12 15 10

المطلوب: إيجاد الوسط الحسابي لهذه النتائج.

الجواب: بتطبيق المعادلة 1-4 أو 2-4، يكون :

$$\bar{x} = \frac{10 + 15 + 12 + 13 + 12 + 10 + 12 + 14 + 13 + 12}{10} = 12.3$$

إذن الوسط الحسابي لنتائج الطالب، أي معدل هو : 12.3

2- الوسط الحسابي للبيانات المبوبة : كما سبق وأن رأينا في الفصل الثاني فإن البيانات ذات الصفات الكمية المبوبة تكون على نوعين، إما مدى فئاتها معدوم أي : $L = 0$ ، أو مدى فئاتها أكبر من الصفر أي : $L > 0$

1-الوسط الحسابي للبيانات المبوبة التي مدى فئاتها معدوم : في هذه الحالة، بدل جمع القيمة الواحدة x_i عدة مرات حسب موضعها بين القيم، فإنه يتم ضربها في عدد تكراراتها التي نرمز لها بـ f_i وجمع النواتج وقسمتها على مجموع التكرارات، لأن التكرار يعبر عن عدد مرات تكرار القيمة x_i ضمن مجموعة القيم، وتكون بذلك كل قيمة مرجحة بتكرارها، لذلك يسمى الوسط الحسابي في هذه الحالة أحيانا بالوسط الحسابي المرجح.

تعريف 4-2: إذا كانت لدينا البيانات : $x_1, x_2, x_3, \dots, x_n$

تكراراتها على التوالي : $f_1, f_2, f_3, \dots, f_n$

فإن وسطها الحسابي يعطى بالصيغة التالية :

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_n f_n}{f_1 + f_2 + f_3 + \dots + f_n} \quad 3-4$$

وإختصارا يكتب بالعارة التالية :

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i} \quad 4-4$$

حيث : $i=1,2,3,\dots,n$ (عدد الفئات).

مثال 4-2: بوب بيانات المثال 4-1 في جدول تكراري غير مستمر (طول فئاته : $L = 0$)، ثم أوجد الوسط الحسابي من الجدول المحصل عليه، وقارنه بالوسط الحسابي للبيانات الأصلية كما وردت في المثال 4-1.

بتطبيق قواعد التبويب كما وردت في الفصل الثاني نحصل على الجدول التالي:

f_i	x_i	i
2	10	1
4	12	2
2	13	3
1	14	4
1	15	5
10	/	مج

جدول 1-4

لإيجاد الوسط الحسابي لبيانات الجدول المحصل عليه، نطبق المعادلة رقم 3-4 أو 4-4، ولأجل ذلك نضيف عموداً إلى الجدول السابق، لحساب جداءات القيم في تكراراتها، أي: $x_i f_i$ ، كما هو واضح في الجدول 2-4 أدناه، ثم نجمع النواتج، ونقسمها على مجموع التكرارات.

$x_i f_i$	f_i	x_i	i
20	2	10	1
48	4	12	2
26	2	13	3
14	1	14	4
15	1	15	5
123	10	/	مج

جدول 2-4

من الجدول 2-4، نجد أن: $\sum_{i=1}^5 x_i f_i = 123$ و $\sum_{i=1}^5 f_i = 10$

وبتطبيق المعادلة رقم 4-4 يكون: $\bar{x} = \frac{\sum_{i=1}^5 x_i f_i}{\sum_{i=1}^5 f_i} = \frac{123}{10} = 12.3$

وهي نفس النتيجة المحصل عليها في المثال 1-4.

جـ- الوسط الحسابي للبيانات المبوبة التي طول فئاتها أكبر من الصفر: في هذه الحالة يستحيل إيجاد الوسط الحسابي الحقيقي، كما في حالة البيانات غير المبوبة أو

البيانات المبوبة التي مدى فئاتها معدوم، لأن مدى الفئات قد لا يكون متساو بالنسبة لجميع الفئات، كما أن توزيع القيم داخل هذا المجال قد يكون غير متماثل، لذلك يتم إيجاد وسط تقريبي بالإعتماد على مراكز الفئات التي سوف نرمز لها بـ c_i ، وذلك باستخدام القاعدة التالية:

$$\bar{x} = \frac{\sum_{i=1}^n c_i f_i}{\sum_{i=1}^n f_i} \quad 5-4$$

حيث: c_i مركز الفئة i (أنظر المعادلة 3-4)

مثال 3-4: البيانات التالية تمثل الأجور الأسبوعية التي يتقاضاها عمال أحد المصانع بآلاف الدينارات:

18	19	15	14	10	12	11	10
24	21	23	20	17	16	15	15
27	25	22	24	23	20	24	23
25	29	28	27	25	25	29	28
32	31	30	26	27	25	27	28
34	32	34	30	30	33	34	33

المطلوب:

- 1- أوجد الوسط الحسابي للأجور.
- 2- بوب البيانات أعلاه في جدول تكراري مستمر طول فئاته: $L=5.10^3$ دينار.
- 3- من البيانات المحصل عليها من السؤال 2 أوجد الوسط الحسابي وقارنه بالوسط المحصل عليه من السؤال 1.

الإجابة:

- 1- إيجاد الوسط الحسابي : يعطى الوسط الحسابي للبيانات غير المبوبة بالمعادلة 2-4، أي:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N}$$

مجموع القيم هو : 1150، بينما عدد القيم هو : 48 ومنه نجد:

$$\bar{x} = \frac{1150}{48} = 23.95$$

و يعني هذا أن متوسط أجور العمال هو : 23.95. 10 دينار.

2- تبويب البيانات في جدول تكراري طول فئاته 5. 10 دينار: باستخدام مبادئ التبويب الواردة في الفصل الثاني نحصل على التوزيع التالي :

f_i	الفئات	i
5	10-15	1
7	15-20	2
10	20-25	3
15	25-30	4
11	30-35	5
48		مج

جدول 3-4

3- إيجاد الوسط الحسابي لبيانات الجدول 3-4 المحصل عليه بعد التبويب: التوزيع الجديد هو توزيع تكراري مستمر لذلك يتم تطبيق العلاقة 4-5 لإيجاد الوسط الحسابي وهي :

$$\bar{x} = \frac{\sum_{i=1}^5 c_i f_i}{\sum_{i=1}^5 f_i}$$

ولأجل تطبيق هذه العلاقة تتم إضافة عمودين للجدول السابق أحدهما نحسب فيه مراكز الفئات والآخر نحسب فيه مراكز الفئات مضروبة في التكرارات، وهو ما يوضحه الجدول 4-4.

i	الفئات	f _i	C _i	C _i f _i
1	10-15	5	12.5	62.5
2	15-20	7	17.5	122.5
3	20-25	10	22.5	225.0
4	25-30	15	27.5	412.5
5	30-35	11	32.5	357.5
مج		48		1180.0

جدول 4-4

من الجدول نستنتج أن قيمة بسط المعادلة المشار إليها هو: 1180، ومقامها هو: 48 ومنه يكون :

$$\bar{x} = \frac{1180.0}{48} = 24.58$$

أي أن الوسط الحسابي للبيانات الجديدة هو: 24.58 . 10³ دينار.

واضح تماما بأن الوسط الحسابي للبيانات قبل وضعها في جدول مستمر، يختلف قليلا عنها لما وضعت في جدول تكراري مستمر، لأنه في الحالة الأخيرة يعتمد الوسط الحسابي فقط على الحدود العليا والدنيا للفئات وليس على جميع القيم الأصلية.

3- خواص الوسط الحسابي : يتميز الوسط الحسابي بمجموعة من الخواص تجعله أفضل مقياس من مقاييس الترتبة المركزية، في المعالجة الإحصائية، منها ما يلي :

1- الخاصية الأولى : المجموع الجبري لفروقات القيم عن وسطها الحسابي، يكون دائما معدوما.
ففي حالة البيانات غير المبوبة يكون :

$$\sum_{i=1}^n (x_i - \bar{x}) = 0 \quad 6-4$$

أما في حالة البيانات المبوبة فيكون :

$$\sum_{i=1}^n (x_i - \bar{x}) f_i = 0 \quad 7-4$$

الاثبات : يتم إثبات هذه الخاصية رياضيا و حسابيا كما يلي :

* الإثبات الرياضي :

** حالة البيانات غير المبوبة : بنشر المعادلة رقم 6-4،

نحصل على مايلي :

$$\sum_{i=1}^n (x_i - \bar{x}) = (x_1 - \bar{x}) + (x_2 - \bar{x}) + (x_3 - \bar{x}) + \dots + (x_n - \bar{x})$$

بفك الأقواس و جمع الحدود المتشابهة نجد :

$$\sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - N\bar{x} \quad 8-4$$

ومعلوم لدينا أن الوسط الحسابي للبيانات غير المبوبة يعطى كما يلي :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N}$$

بضرب طرفي العلاقة في وسطها نجد :

$$\sum_{i=1}^n x_i = N\bar{x} \quad 9-4$$

بتعويض العلاقة 9-4 في العلاقة 8-4 نجد :

$$\sum_{i=1}^n (x_i - \bar{x}) = N\bar{x} - N\bar{x} = 0$$

و بالتالي فان الخاصية صحيحة في حالة البيانات غير المبوبة.

** حالة البيانات المبوبة :

بنشر المعادلة 4-7 نجد :

$$\sum_{i=1}^n (x_i - \bar{x})f_i = (x_1 - \bar{x})f_1 + (x_2 - \bar{x})f_2 + (x_3 - \bar{x})f_3 + \dots + (x_n - \bar{x})f_n$$

$$\sum_{i=1}^n (x_i - \bar{x})f_i = \sum_{i=1}^n x_i f_i - \bar{x} \sum_{i=1}^n f_i \quad \text{أي: 4-10}$$

معلوم لدينا أن الوسط الحسابي للبيانات المبوبة التي طول فئاتها معدوم هو:

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i}$$

بضرب الطرفين في الوسطين نجد :

$$\bar{x} \sum_{i=1}^n f_i = \sum_{i=1}^n x_i f_i \quad 4-11$$

بتعويض 4-11 في 4-10 نجد :

$$\sum_{i=1}^n (x_i - \bar{x})f_i = \bar{x} \sum_{i=1}^n f_i - \bar{x} \sum_{i=1}^n f_i = 0$$

وبالتالي فإن الخاصية أيضا صحيحة في حالة البيانات المبوبة.

***الاثبات العكسي :** ويتم عن طريق المثالين 4-4 و 4-5 التاليين :

مثال 4-4 : من بيانات المثال رقم 4-1، أثبت صحة المعادلة رقم 4-6 حسابيا.

الاثبات : البيانات المشار إليها هي :

10 15 12 13 12 10 12 14 13 12
وسطها الحسابي هو:

$$\bar{x} = \frac{123}{10} = 12.3$$

بتطبيق العلاقة 4-6 المشار إليها أعلاه من خلال الجدول التالي، نجد بالفعل أن مجموع الانحرافات (الفروقات) عن الوسط الحسابي معدوم، وذلك ما يوضحه الجدول 4-5.

x_i	10	15	12	13	12	10	12	14	13	12	مج
$x_i - 12.3$	-2.3	2.7	-0.3	0.7	-0.3	-2.3	-0.3	1.7	0.7	-0.3	0.0

جدول 4-5

من السطر الأخير في الجدول يظهر أن مجموع الانحرافات عن الوسط الحسابي معدوم.

مثال 4-5: أثبت صحة العلاقة رقم 4-7 من خلال بيانات الجدول 4-1.

الإجابة: الوسط الحسابي للبيانات المشار إليها هو :

$$\bar{x} = \frac{123}{10} = 12.3$$

بتطبيق العلاقة 4-7 المشار إليها نجدها صحيحة وذلك من خلال الجدول التالي :

i	x_i	f_i	$(x_i - \bar{x})$	$(x_i - \bar{x})f_i$
1	10	2	-2.3	-4.6
2	12	4	-0.3	-1.2
3	13	2	0.7	1.4
4	14	1	1.7	1.7
5	15	1	2.7	2.7
مج	/	10		0.0

جدول 4-6

اذ أن مجموع عناصر العمود الأخير تساوي الصفر أي :

$$\sum_{i=1}^n (x_i - \bar{x})f_i = 0$$

ب- الخاصية الثانية : إذا كانت لدينا مجموعة من القيم وسطها الحسابي :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} \quad \text{في حالة البيانات غير المبوبة}$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i} \quad \text{أو : في حالة البيانات المبوبة}$$

وأضيفت قيمة ثابتة c الى كل قيمة من القيم الأصلية، فإن الوسط الحسابي للقيم الجديدة يساوي الى الوسط الحسابي للقيم الأصلية مضافا إليه القيمة الثابتة، أي :

$$\bar{x} = \bar{x} + c \quad 12-4$$

حيث \bar{x} : الوسط الحسابي للقيم الجديدة.
هذه الخاصية صحيحة أيضا في حالة طرح المقدار c من القيم الأصلية، حيث أن الوسط الحسابي للقيم الجديدة يعطى بالعلاقة التالية :

$$\bar{x} = \bar{x} - c \quad 13-4$$

الاثبات : يتم اثبات هذه الخاصية أيضا اثباتا رياضيا واثباتا حسابيا كما يلي:

• الإثبات الرياضي :

** حالة البيانات غير المبوبة : إذا كانت لدينا البيانات : $x_1, x_2, x_3, \dots, x_n$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} \quad \text{فان وسطها الحسابي هو:}$$

عند إضافة القيمة الثابتة c الى كل قيمة من تلك القيم تصبح القيم الجديدة كما يلي:

$$x_1 + c, x_2 + c, x_3 + c, \dots, x_n + c$$

و يكون الوسط الحسابي للقيم الجديدة هذه على النحو:

$$\bar{x} = \frac{x_1 + c + x_2 + c + x_3 + c + \dots + x_n + c}{N} \quad 14-4$$

وهو ما يمكن كتابته كما يلي :

$$\bar{x} = \frac{\sum_{i=1}^n x_i + Nc}{N}$$

15-4

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} + c$$

16-4

أو :

ومعلوم أن القسم الأول في الطرف الأيمن هو الوسط الحسابي، للبيانات الأصلية، وبالتالي فالعلاقة رقم 12-4 صحيحة.

•• حالة البيانات المبوبة: عند البرهان على هذه الخاصية يجب أخذ بعين الاعتبار تكرارات كل قيمة، فإذا كانت لدينا البيانات:

$x_1, x_2, x_3, \dots, x_n$ تكراراتها على التوالي: $f_1, f_2, f_3, \dots, f_n$ فإن وسطها الحسابي يعطى بالمعادلة 4-4، أي :

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i}$$

عند إضافة المقدار c الى كل قيمة من قيم هذه البيانات نحصل على الجدول التالي:

i	$x_i + c$	f_i
1	$x_1 + c$	f_1
2	$x_2 + c$	f_2
3	$x_3 + c$	f_3
.	.	.
.	.	.
.	.	.
N	$x_n + c$	f_n

جدول 4-7

وطبعا يكون الوسط الحسابي لهذه البيانات على شكل المعادلة التالية :

$$\bar{x} = \frac{\sum_{i=1}^n (x_i + c)f_i}{\sum_{i=1}^n f_i} \quad 17-4$$

بفك بسط المعادلة 17-4 نجد :

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i + c \sum_{i=1}^n f_i}{\sum_{i=1}^n f_i} \quad 18-4$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i} + c \quad 19-4 \quad \text{ومنه نجد :}$$

معلوم أن القسم الأول من الطرف الأيمن للمعادلة 19-4 هو الوسط الحسابي للبيانات الأصلية، ومنه فإن الخاصية صحيحة أيضا في حالة البيانات المبوبة.

•• الإثبات الحاسبي : يتم ذلك عن طريق المثالين التاليين :

مثال 4-6 : من بيانات المثال 4-1 اثبت أنه إذا طرح المقدار $c=2$ من كل قيمة من القيم فإن الوسط الحسابي للقيم الجديدة، يساوي الى الوسط الحسابي للقيم الأصلية وهو 12.3 منقوصا منه 2، أي الوسط الحسابي الجديد هو : 10.3.

ب طرح المقدار $C=2$ من بيانات المثال المشار اليه نحصل على القيم التالية:

8 13 10 11 10 8 10 12 11 10

و يكون وسطها الحسابي هو :

$$\bar{x} = \frac{103}{10} = 10.3$$

أي : $\bar{x} = 12.3 - 2$ و بالتالي فإن الخاصية صحيحة.

مثال 4-7 : من بيانات الجدول 4-1، أثبت أنه إذا طرح المقدار $c=2$ من كل قيمة من القيم فإن الوسط الحسابي للقيم الأصلية يساوي الى : $10.2=12.3-2$ ، أي الوسط الحساب للقيم الأصلية منقوصاً منه: $c=2$

الإجابة: عند طرح المقدار $c=2$ من البيانات المشار إليها نحصل على بيانات العمود الرابع من الجدول 4-8، و بإجراء الحسابات من خلال الجزء الأيمن من الجدول 4-8 أدناه نجد :

i	x_i	f_i	(x_i-2)	$(x_i-2)f_i$
1	10	2	8	16
2	12	4	10	40
3	13	2	11	22
4	14	1	12	12
5	15	1	13	13
مج		10		103

جدول 4-8

الوسط الحسابي للبيانات الجديدة هو :

$$\bar{x} = \frac{103}{10} = 10.3$$

$$\bar{x} = 12.3 - 2 = 10.3 \quad \text{أي :}$$

وبالتالي فإن الخاصية صحيحة.

ج-الخاصية الثالثة : إذا كانت لدينا مجموعة من القيم وسطها الحسابي هو:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} \quad \text{في حالة البيانات غير المبوبة}$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i} \quad \text{في حالة البيانات المبوبة} \quad \text{أو :}$$

وضربت كل قيمة من تلك القيم في عدد ثابت وليكن :c،
فإن الوسط الحسابي للقيم الجديدة، يساوي الى الوسط الحسابي
للقيم الأصلية مضروباً في العدد c، أي:

$$\bar{x} = c \cdot \bar{x} \quad 20-4$$

الاثبات : يتم إثبات الخاصية رياضياً وحسابياً كما يلي :

***الاثبات الرياضي :**

****حالة البيانات غير المبوبة :** إذا كانت مجموعة القيم هي:

$x_1, x_2, x_3, \dots, x_n$ فإن وسطها الحسابي هو :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N}$$

بضرب هذه القيم في مقدار ثابت C ، نحصل على القيم
الجديدة :

$c \cdot x_1, c \cdot x_2, c \cdot x_3, \dots, c \cdot x_n$

ويكون الوسط الحسابي لهذه القيم هو:

$$\bar{x} = \frac{\sum_{i=1}^n (c \cdot x_i)}{N} \quad 21-4$$

بإخراج C عامل مشترك، أي خارج المجموع، نحصل على المعادلة :

$$\bar{x} = c \cdot \frac{\sum_{i=1}^n x_i}{N} \quad 22-4$$

معلوم أن الطرف الأخير من المعادلة 4-22 هو الوسط الحسابي للقيم الأصلية، ومنه نستنتج أن العلاقة 4-20 صحيحة، وبالتالي فإن الخاصية صحيحة.

•• حالة البيانات المبوبة: لإثبات الخاصية في حالة البيانات المبوبة يجب أخذ التكرارات بعين الاعتبار، فإذا كانت لدينا البيانات:

$x_1, x_2, x_3, \dots, x_n$ تكراراتها على التوالي:

$f_1, f_2, f_3, \dots, f_n$ فإن وسطها الحسابي هو:

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i}$$

بضربها في المقدار c نحصل على البيانات الجديدة كما هي في الجدول التالي:

i	$c.x_i$	f_i
1	$c.x_1$	f_1
2	$c.x_2$	f_2
3	$c.x_3$	f_3
.	.	.
.	.	.
.	.	.
N	$c.x_n$	f_n

جدول 4-9

و يكون الوسط الحسابي لهذه البيانات على النحو:

$$\bar{x} = \frac{\sum_{i=1}^n (c.x_i) f_i}{\sum_{i=1}^n f_i}$$

باخراج C عامل مشترك، أي خارج المجموع نجد:

$$\bar{x} = c \cdot \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i} \quad 23-4$$

الطرف الأخير من المعادلة 23-4 ، هو الوسط الحسابي ، وبالتالي نستنتج أن العلاقة 20-4 صحيحة ، أي أن الخاصية صحيحة أيضا في حالة البيانات المبوبة.

تتم البرهنة بنفس الطريقة في حالة قسمة البيانات الأصلية على عدد ثابت c ، حيث يساوي الوسط الحسابي الجديد الى الوسط الحسابي للبيانات الأصلية مقسوما على العدد c .

• **الاثبات العملي :** يتم كذلك عن طريق المثالين التاليين :

مثال 4-8: اثبت أنه إذا ضربت بيانات المثال 4-1 في المقدار $c=2$ ، فإن الوسط

الحسابي للبيانات الجديدة هو: $\bar{x} = 2 \cdot \bar{x} = 2 \times 12.3 = 24.6$

بضرب البيانات المشار إليها في العدد $c=2$ ، نحصل على القيم التالية:

20 30 24 26 24 20 24 28 26 24

ويكون الوسط الحسابي لهذه البيانات هو: $\bar{x} = \frac{246}{10} = 24.6$

أي أن الوسط الحسابي الجديد يساوي الى الوسط الحسابي للبيانات الأصلية مضروبا في: $c=2$

مثال 4-9: اثبت أنه إذا ضربت بيانات الجدول 4-1 في المقدار $c=2$ فإن

الوسط الحسابي للبيانات الجديدة هو: $\bar{x} = 2 \cdot \bar{x} = 2 \cdot 12.3 = 24.6$

الإجابة: لإثبات ذلك نجري الحسابات الضرورية في الجدول التالي :

i	x_i	f_i	$2 \cdot x_i$	$2 \cdot x_i \cdot f_i$
1	10	2	20	40
2	12	4	24	96
3	13	2	26	52
4	14	1	28	28
5	15	1	30	30
مج		10		246

جدول 4-10

بتطبيق القاعدة يكون الوسط الحسابي الجديد كما يلي :

$$\bar{x} = \frac{246}{10} = 24.6$$

$$\bar{x} = 2 \cdot \bar{x} = 2 \cdot 12.3 = 24.6 \quad \text{أي :}$$

و بالتالي فإن الخاصية صحيحة.

د- الخاصية الرابعة : مجموع مربعات فروقات القيم عن وسطها الحسابي، يكون دائما أقل من مجموع مربعات فروقات تلك القيم عن أية قيمة أخرى، مهما كانت تختلف عن الوسط الحسابي، أي:

$$\sum_{i=1}^n (x_i - \bar{x})^2 < \sum_{i=1}^n (x_i - c)^2, \quad \forall \bar{x} \neq c \quad 24-4$$

هذا في حالة البيانات غير المبوبة أما في حالة البيانات المبوبة فيكون:

$$\sum_{i=1}^n (x_i - \bar{x})^2 f_i < \sum_{i=1}^n (x_i - c)^2 f_i, \quad \forall \bar{x} \neq c \quad 25-4$$

الإثبات :

• **حالة البيانات غير المبوبة :** تعني العبارة 24-4 أن الطرف الأيسر منها يكون دائما في قيمته الدنيا أي مجموع مربعات فروقات القيم عن الوسط الحسابي يكون دائما في أدنى قيمة، و يعني ذلك أيضا أن الوسط الحسابي لمربعات تلك القيم يكون في أدنى قيمة، أي أن العبارة التالية تكون في أدنى قيمة لها

$$K = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N}$$

لإثبات ذلك نستبدل الوسط الحسابي في المعادلة أعلاه بقيمة أخرى و لتكن Z، ثم نثبت أن العبارة K لاتأخذ قيمتها الدنيا الا إذا كانت القيمة Z تساوي الوسط الحسابي.

$$K = \frac{\sum_{i=1}^n (x_i - Z)^2}{N}$$

26-4

من المعلوم أن K تكون في أدنى قيمة لها إذا توافر شرطان، الأول و هو الشرط اللازم و هو أن تكون مشتقتها الأولى معدومة، الثاني هو الشرط الكافي و فيه يجب أن تكون مشتقتها الثانية أكبر من الصفر، لذلك نوجد أولاً المشتقة الأولى ونساويها الى الصفر ونوجد من خلال ذلك القيمة Z التي تجعل K معدومة، ثم نرى إذا ما كانت المشتقة الثانية أكبر من الصفر.

$$\frac{\delta K}{\delta Z} = \frac{-2 \sum_{i=1}^n (x_i - Z)}{N} = 0$$

ومنه يكون:

$$\frac{\delta K}{\delta Z} = \frac{\sum_{i=1}^n (x_i - Z)}{N} = 0$$

$$\frac{\delta K}{\delta Z} = \frac{\sum_{i=1}^n x_i}{N} - \frac{N \cdot Z}{N} = 0$$

بفك القوس نجد :

الطرف الأول هو الوسط الحسابي، ومنه نجد : $x - Z = 0$ أي :

$$Z = \bar{x}$$

بما أن المشتقة الثانية للمعادلة 4-26 أعلاه أكبر من الصفر، لذلك فإن القيمة الوحيدة التي تجعل المعادلة 4-26 في أدنى قيمة لها هي الوسط الحسابي، و بالتالي فإن العبارة 4-24 صحيحة.

*** حالة البيانات المبوبة:** العبارة 4-25 تعني أن طرفها الأيسر يكون دائما في أدنى قيمة، و يعني ذلك أيضا أن الوسط الحسابي لهذا الطرف يكون في أدنى قيمة، بمعنى أن الوسط الحسابي لمربعات الفروقات يكون في أدنى قيمة له. إذا استبدلنا الوسط الحسابي بالقيمة Z للطرف المشار اليه علينا أن نثبت بأن العبارة 4-27 :

$$K = \frac{\sum_{i=1}^n (x_i - Z)^2 f_i}{\sum_{i=1}^n f_i} \quad 27-4$$

تكون في أدنى قيمة لها عندما تكون قيمة Z تساوي الوسط الحسابي، وذلك إذا ما توافر شرطا النهاية الصغرى. بالإشتقاق بالنسبة الى Z ، ومساواة النتيجة الى الصفر نجد :

$$\frac{\delta K}{\delta Z} = \frac{-2 \sum_{i=1}^n (x_i - Z) f_i}{\sum_{i=1}^n f_i} = 0$$

$$\frac{\delta K}{\delta Z} = \frac{\sum_{i=1}^n (x_i - Z)f_i}{\sum_{i=1}^n f_i} = 0 \quad \text{و هذا يكفي:}$$

$$\frac{\delta K}{\delta Z} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i} - \frac{Z \sum_{i=1}^n f_i}{\sum_{i=1}^n f_i} = 0 \quad \text{بفك القوس نجد:}$$

الطرف الأول هو الوسط الحسابي، ومنه نجد: $\bar{x} - Z = 0$ أي: $Z = \bar{x}$.
 بما أن المشتقة الثانية للمعادلة 4-27 أعلاه أكبر من الصفر،
 لذلك فإن القيمة الوحيدة التي تجعل المعادلة 4-27 في أدنى قيمة
 لها هي الوسط الحسابي، و بالتالي فإن العبارة 4-25 صحيحة.
 يمكن اثبات هذه الخاصية أيضا بطريقة أخرى، نعطي فكرتها
 فقط على حالة البيانات غير المبوبة، وذلك كما يلي:
 بفرض أنه لدينا مجموعة من القيم، فإن إنحرافها عن قيمة
 ما ولتكن: Z هو:

$$(x_i - Z) = (x_i - \bar{x}) - (Z - \bar{x})$$

بتربيع الطرفين نحصل:

$$(x_i - Z)^2 = (x_i - \bar{x})^2 - 2(x_i - \bar{x})(Z - \bar{x}) + (Z - \bar{x})^2$$

بإدخال المجموع نحصل:

$$\sum_{i=1}^n (x_i - Z)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 - 2(Z - \bar{x}) \sum_{i=1}^n (x_i - \bar{x}) + N(Z - \bar{x})^2$$

$$\sum_{i=1}^n (x_i - \bar{x}) = 0 \quad \text{بالإستفادة من الخاصية الأولى للوسط الحسابي وهي:}$$

$$\sum_{i=1}^n (x_i - Z)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + N(Z - \bar{x})^2 \quad \text{نجد: 28-4}$$

نلاحظ أن الطرف الثاني يكون في أدنى قيمة له عندما يكون :

$$N(Z - \bar{x})^2 = 0$$

و يتحقق ذلك فقط عندما يكون Z يساوي الوسط الحسابي.
وفي حالة البيانات المبوبة تتم البرهنة بنفس الطريقة أو بضرب المعادلة 28-4 في f_i :

مثال 4-10 : أثبت حسابيا أن مجموع مربعات فروقات بيانات الجدول 4-1، عن وسطها الحسابي، أصغر من مجموع مربعات فروقات هذه القيم عن القيمة : $c = 20$ وعن القيمة : $c = -5$

لإثبات ذلك نجري الحسابات الضرورية من خلال الجدول التالي :

i	x_i	f_i	$(x_i - \bar{x})^2 f_i$	$(x_i - 20)^2 f_i$	$(x_i - (-5))^2 f_i$
1	10	2	10.58	200	450
2	12	4	0.36	256	1156
3	13	2	0.98	98	648
4	14	1	2.89	36	361
5	15	1	7.29	25	400
مج		10	22.10	615	3015

جدول 4-11

من الجدول نلاحظ ما يلي :

$$\sum_{i=1}^n (x_i - \bar{x})^2 f_i = \sum_{i=1}^n (x_i - 12.3)^2 f_i = 22.10$$

$$\sum_{i=1}^n (x_i - 20)^2 f_i = 615$$

$$\sum_{i=1}^n (x_i - (-5))^2 f_i = 3015$$

أي أن مربع الفروقات عن الوسط الحسابي هو الأصغر، وبالتالي فإن الخاصية صحيحة حسابيا.

هـ-الخاصية الخامسة : يمكن إيجاد الوسط الحسابي لأي مجموعة من البيانات باستخدام وسط فرضي، عن طريق المعادلة التالية:

$$x = Z + \frac{\sum_{i=1}^n (x_i - Z)}{N} \quad 29-4$$

هذا في حالة البيانات غير المبوبة، أما في حالة البيانات المبوبة، فيتم ذلك عن طريق المعادلة :

$$\bar{x} = Z + \frac{\sum_{i=1}^n (x_i - Z)f_i}{\sum_{i=1}^n f_i} \quad 30-4$$

حيث Z : وسط فرضي تختلف قيمته عن الوسط الحسابي، أما إذا ساوت قيمته قيمة الوسط الحسابي، فإن الطرف الأخير في كل من المعادلتين 29-4 و 30-4 يساوي الصفر.

و تسمى طريقة إيجاد الوسط الحسابي باستخدام الوسط الفرضي، بطريقة الفروقات، ويمكن إثبات صحة المعادلتين 29 و 30-4 بنشرهما اذ نجد بأن الطرفين متساويين.

مثال 4-11: أوجد الوسط الحسابي لبيانات الجدول 4-1، باستخدام الوسط الفرضي: $Z=20$.

بما أن البيانات المشار إليها هي بيانات مبوبة، لذلك نطبق المعادلة 30-4، ونجري الحسابات الضرورية من خلال الجدول التالي :

i	x_i	f_i	$(x_i - 20)f_i$
1	10	2	-20
2	12	4	-32
3	13	2	-14
4	14	1	-6
5	15	1	-5
مج		10	-77

جدول 4-12

من الجدول 4-12 نجد: $\sum_{i=1}^n (x_i - 20)f_i = -77$

ومنه يكون: $\bar{x} = 20 + \frac{-77}{10} = 12.3$

وهي نفس قيمة الوسط الحسابي المحصل عليه في الأمثلة السابقة.

و- الخاصية السادسة: إذا كانت لدينا القيم:

وأضيفت إلى كل قيمة منها القيم: $x_1, x_2, x_3, \dots, x_n$

على التوالي، فإن الوسط الحسابي للقيم الجديدة: $y_1, y_2, y_3, \dots, y_n$

يكتب على النحو: $z_1, z_2, z_3, \dots, z_n$

$$\bar{z} = \bar{x} + \bar{y} \quad 31-4$$

حيث: \bar{x} : الوسط الحسابي للقيم الأصلية. \bar{y} : الوسط الحسابي للقيم

المضافة. \bar{z} : الوسط الحسابي للقيم الجديدة.

ويمكن البرهنة أيضاً على هذه الخاصية سواء في حالة البيانات

غير المبوبة أو في حالة البيانات المبوبة، كما يلي:

الاثبات:

• حالة البيانات غير المبوبة:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N}$$

الوسط الحسابي للبيانات الأصلية هو:

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{N}$$

الوسط الحسابي للبيانات المضافة هو :

عند إضافة مجموعة القيم الى القيم الأصلية، يكون الوسط الحسابي للقيم الجديدة على النحو:

$$\bar{Z} = \frac{\sum_{i=1}^n (x_i + y_i)}{N}$$

32-4

وبفك البسط يكون :

$$\bar{Z} = \frac{\sum_{i=1}^n x_i + \sum_{i=1}^n y_i}{N} = \frac{\sum_{i=1}^n x_i}{N} + \frac{\sum_{i=1}^n y_i}{N}$$

بتعويض الطرفين الأخيرين بقيمهما كما هي أعلاه نجد العلاقة 31-4 محققة.

• **حالة البيانات المبوبة :** في هذه الحالة يكون الوسط الحسابي للبيانات الأصلية كما يلي:

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i}$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i f_i}{\sum_{i=1}^n f_i}$$

أما الوسط الحسابي للبيانات المضافة فهو:

ويكون الوسط الحسابي للبيانات الجديدة كما يلي :

$$\bar{Z} = \frac{\sum_{i=1}^n (x_i + y_i) f_i}{\sum_{i=1}^n f_i}$$

33-4

$$\bar{Z} = \frac{\sum_{i=1}^n x_i f_i + \sum_{i=1}^n y_i f_i}{\sum_{i=1}^n f_i} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i} + \frac{\sum_{i=1}^n y_i f_i}{\sum_{i=1}^n f_i}$$

بفك البسط نجد:

بتعويض الطرفين الآخرين بقيمهما نجد أن العلاقة 31-4 صحيحة.

هذه هي أهم خصائص الوسط الحسابي، وهي خصائص يستعان بها إما في البرهنة على صحة الوسط الحسابي كما في الخاصية الأولى أو في تبسيط الحسابات خاصة عندما تكون المعطيات الرقمية كبيرة جداً، و قد لاحظنا أنه يحقق كل شروط يول وهذا ما يجعله من أهم مقاييس الترتيب المركزية وأكثرها استخداماً في كافة مجالات التحليل الكمي للظواهر الاقتصادية والاجتماعية وغيرها، غير أن أهم عيوبه هي أنه يتأثر بالقيم المتطرفة إذ ينجذب إليها وذلك لوزنها المؤثر ضمن مجموعة القيم، ومن عيوبه أيضاً أنه في حالة التوزيعات التكرارية المستمرة المغلقة، لا يمكن إيجاد الوسط الحسابي بكل دقة كما تمت البرهنة على ذلك خلال هذا الفصل، كما أنه في حالة التوزيعات التكرارية المستمرة المفتوحة يستحيل إيجاد الوسط الحسابي لعدم إمكانية إيجاد مركز الفئة الأولى أو مركز الفئة الأخيرة أو الاثنان معاً، غير أن الباحث الإحصائي إذا ما كان ذو حنكة وتجربة فإنه يستطيع تقدير الحد الأدنى لأول فئة أو الحد الأعلى لآخر فئة، اعتماداً على توزيع بقية الفئات، وبالتالي يمكنه إيجاد وسط حسابي تقريبي للبيانات محل الدراسة.

ثانياً: الوسيط: الوسيط هو ثاني مقياس من مقاييس الترتيب المركزية، الأكثر شيوعاً، ويعرف كما يلي :

تعريفه 3-4: وسيط أية مجموعة من القيم، هو القيمة التي تقع في وسط تلك القيم بعد ترتيبها تصاعدياً أو تنازلياً، وتختلف طرق حسابه باختلاف طبيعة البيانات.

1- وسيط البيانات غير الموزونة: يتم حساب الوسيط لهذه البيانات بإتباع الخطوات التالية :

- أ- نقوم بترتيب البيانات تصاعدياً أو تنازلياً.
- ب- نقوم بحساب ترتيب الوسيط c حسب عدد القيم إذا كان زوجياً أو فردياً.
- إذا كان عدد القيم N فردياً يكون :

$$c = \frac{N+1}{2} \quad 34-4$$

وتكون قيمة الوسيط $Mé$ هي القيمة التي ترتيبها c تصاعدياً أو تنازلياً.

- إذا كان عدد القيم N زوجياً يكون :

$$c = \frac{N}{2} \quad 35-4$$

في هذه الحالة نجد قيمتين للوسيط، الأولى ترتيبها c وهي $Mé_1$ نجدها عند العدد التصاعدي، والثانية هي $Mé_2$ ، ترتيبها أيضاً c عند العدد التنازلي أو ترتيبها $c+1$ عند العدد التصاعدي، وبما أن الوسيط يكون وحيداً، أي يأخذ قيمة واحدة فقط ، لذلك تكون قيمته هي الوسط الحسابي للقيمتين $Mé_1$ و $Mé_2$ ، أي :

$$Mé = \frac{Mé_1 + Mé_2}{2} \quad 36-4$$

مثال 4-12 : أوجد وسيط البيانات التالية :

10 8 2 7 3 5 6

لايجاد وسيط هذه البيانات نقوم بترتيبها تصاعديا كما يلي :

2 3 5 6 7 8 10

بما أن عدد القيم فرديا، لذلك نوجد ترتيب الوسيط عن طريق المعادلة 4-34، ويكون :

$$c = \frac{N+1}{2} = \frac{7+1}{2} = 4$$

وعليه تكون قيمة الوسيط هي التي تقع في الرتبة الرابعة، وبالتالي فان:

$$M_e = 6$$

مثال 4-13: أوجد وسيط البيانات التالية :

11 10 8 5 6 3 12 14

لايجاد الوسيط نقوم أيضا بترتيب القيم تصاعديا على النحو :

14 12 11 10 8 6 5 3

و بما أن عدد القيم زوجي لذلك نطبق المعادلة 4-35، و يكون :

$$c = \frac{N}{2} = \frac{8}{2} = 4$$

ويعني هذا أن الوسيط يوجد في الرتبة الرابعة، غير أنه عند العد التصاعدي نجد:

$$M_{e1} = 8$$

$$M_{e2} = 10$$

وعند العد التنازلي نجد :

وعليه يكون الوسيط :

$$M_e = \frac{M_{e1} + M_{e2}}{2} = \frac{8 + 10}{2} = 9$$

2- وسيط البيانات الموزونة : البيانات الموزونة قد تكون غير مستمرة أي مدى فئاتها معدوم، وقد تكون مستمرة أي مدى فئاتها أكبر من الصفر، ويتم إيجاد الوسيط حسب كل حالة كما يلي :

1- طول الفئات معدوم : في هذه الحالة يتم إيجاد الوسيط كما يلي :

* نوجد ترتيب الوسيط باستخدام إحدى المعادلتين التاليتين :

economics

blogspot

الباحث الاقتصادي

108

economics researcher ate

$$c = \frac{\sum_{i=1}^n f_i + 1}{2} \quad 37-4$$

إذا كان مجموع التكرارات فرديا.

$$c = \frac{\sum_{i=1}^n f_i}{2} \quad 38-4$$

إذا كان مجموع التكرارات زوجيا.

* نوجد التكرار المتجمع الصاعد ونبحث عن مكان ترتيب الوسيط بين التكرارات المتجمعة، فنجده بين تكرارين من التكرارات المتجمعة، وتكون قيمة الوسيط هي القيمة المقابلة للتكرار المتجمع اللاحق لترتيب الوسيط.

إذا ما كانت c تساوي إحدى قيم التكرارات المتجمعة فإن M_e تساوي الفئة المقابلة لها، سواء كان مجموع التكرارات زوجيا أو فرديا.

مثال 4-14: أوجد وسيط البيانات التالية :

i	x_i	f_i
1	10	2
2	12	3
3	14	5
4	16	3
5	18	2
مج		15

جدول 4-13

نوجد التكرار المتجمع الصاعد كما هو وارد في الجدول التالي :

i	x_i	f_i	الحدود العليا	ت.م.ص
1	10	2	-12	2
2	12	3	-14	5
3	14	5	-16	10
4	16	3	-18	13
5	18	2	-20	15
مج		15		

جدول 4-14

ثم نوجد ترتيب الوسيط بتطبيق المعادلة 4-37 فيكون :

$$c = \frac{15+1}{2} = 8$$

ترتيب الوسيط يوجد بين التكراريين المتجمعين : 5 و 10،
لذلك فان الوسيط يساوي الى الفئة المقابلة لـ: 10، وبالتالي
يكون :

$$M_e = 14$$

ب- طول الفئات الخبر من الصف : يتم إيجاد الوسيط في
هذه الحالة، بعدة طرق تعطي نتائج متقاربة في الغالب وهي :
• الطريقة الأولى : يتم إيجاد الوسيط حسب هذه الطريقة
باستخدام المنهجية التالية:

- ا- نوجد التكرار المتجمع الصاعد أو التكرار المتجمع النازل.
- ب- نوجد ترتيب الوسيط باستخدام المعادلة التالية :

$$c = \frac{\sum_{i=1}^n f_i}{2} \quad 39-4$$

ج- نبحث عن مكان ترتيب الوسيط بين التكرارات
المتجمعة، فنجده بين تكرارين من التكرارات المتجمعة أحدهما
سابق له و الآخر لاحق له.

د- نبحث عن الفئة الوسيطة في حدود الفئات التي تحدد
التكرار المتجمع، بحيث يكون الحد الأدنى للفئة الوسيطة هو
الحد المقابل للتكرار المتجمع السابق لترتيب الوسيط، وحدها
الأعلى هو الحد المقابل للتكرار المتجمع اللاحق لترتيب
الوسيط.

هـ- نطبق المعادلة التالية لإيجاد قيمة الوسيط :

$$M_e = d + \frac{c - f_{i-1}^+}{f_{i+1}^+ - f_{i-1}^+} \times L \quad 40-4$$

حيث :

M é	قيمة الوسيط.
d	الحد الأدنى للفئة الوسيطة.
c	ترتيب الوسيط.
L	طول الفئة الوسيطة.
f_{i-1}^+	التكرار المتجمع السابق لترتيب الوسيط
f_{i+1}^+	التكرار المتجمع اللاحق لترتيب الوسيط

مثال 4-15: أوجد وسيط البيانات التالية :

i	الفئات	f_i
1	أقل من 10	5
2	10-20	8
3	20-30	6
4	30-40	5
5	40-50	4
6	50-60	2
مج		30

جدول 4-15

الإجابة: لإيجاد الوسيط بالطريقة الأولى نطبق المعادلة 4-40، ولأجل ذلك، نوجد التكرار المتجمع الصاعد كما هو وارد في الجدول أدناه، ونوجد ترتيب الوسيط، بتطبيق المعادلة 4-39، حيث نجد: $c=15$

i	الفئات	f_i	الحد الأعلى	ت.م.ص
1	-10	5	-10	5
2	10-20	8	-20	$f_{i-1}^+ \rightarrow 13$
3	20-30	6	-30	$f_{i+1}^+ \rightarrow 19$
4	30-40	5	-40	24
5	40-50	4	-50	28
6	50-60	2	-60	30
مج		30		

جدول 4-16

و بالتالي نجد : c محصور بين التكرارين المتجمعين (13 < c < 19) كما هو مـشار إليه في الجدول أعلاه، وتكون الفئة الوسيطة هي: (20-30)، ومنه يكون : d=20، L=10، و تكون قيمة الوسيط هي:

$$M_e = 20 + \frac{15 - 13}{19 - 13} \times 10 = 23.33$$

• **الطريقة الثانية :** حسب هذه الطريقة يتم استخدام نفس الخطوات أ، ب، ج، د المشار إليها في الطريقة الأولى، مع استبدال معادلة الخطوة الأخيرة بالمعادلة التالية :

$$M_e = d + \frac{c - f_{i-1}^+}{f_i} \times L \quad 41-4$$

حيث :

M e	قيمة الوسيط.
d	الحد الأدنى للفئة الوسيطة.
c	ترتيب الفئة الوسيطة
L	طول الفئة الوسيطة.
f_{i-1}^+	التكرار المتجمع السابق لترتيب الوسيط
f_i	تكرار الفئة الوسيطة

مثال 4-16: أوجد وسيط بيانات المثال 4-15.

بتطبيق المعادلة 41-4 نجد :

$$M_e = 20 + \frac{15 - 13}{6} \times 10 = 23.33$$

وهي نفس القيمة المحصل عليها باستخدام الطريقة السابقة.

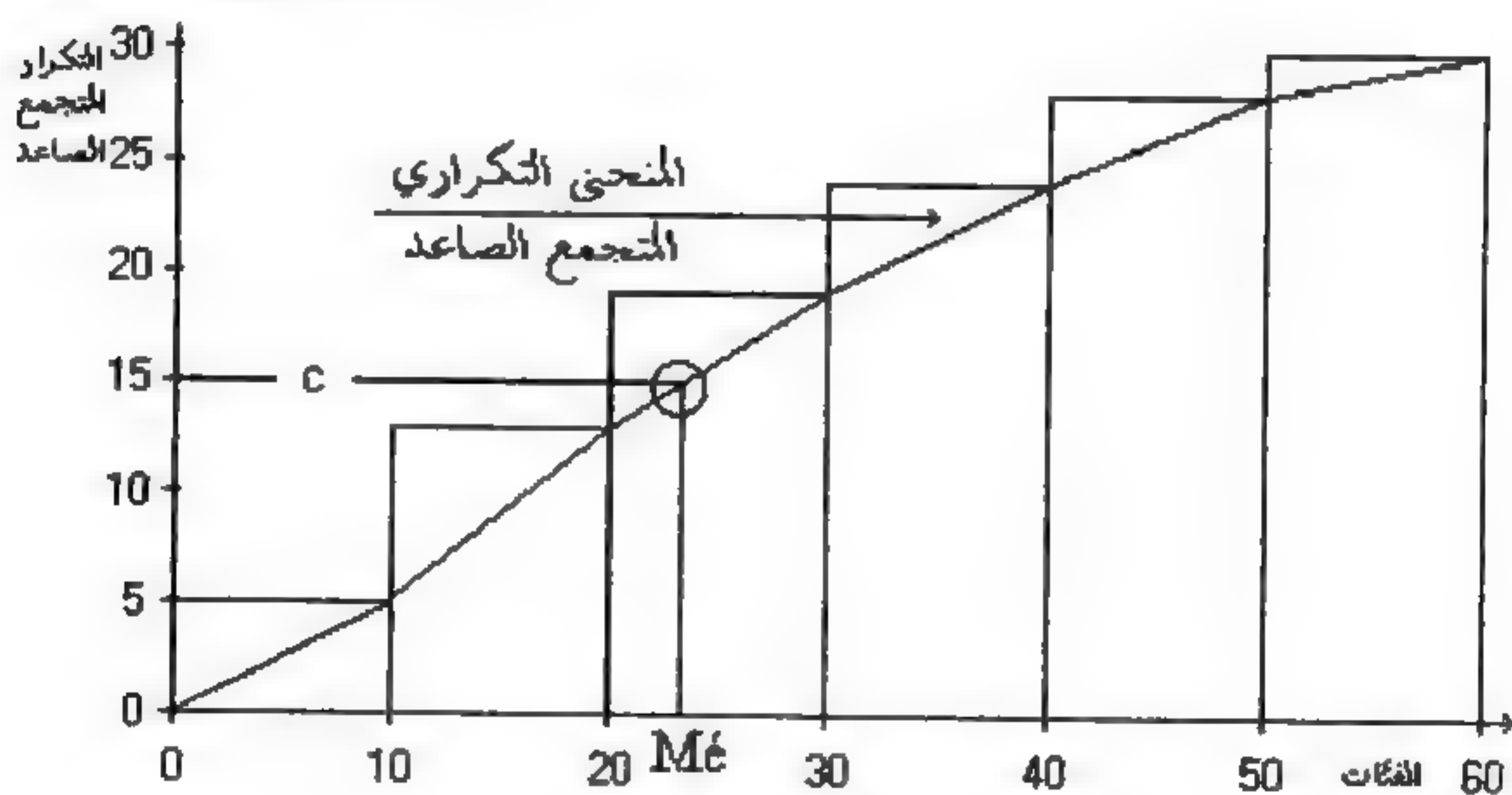
• **الطريقة الثالثة :** وهي الطريقة البيانية، إذ يمكن إيجاد الوسيط بيانياً، وذلك برسم ، إما المنحنى التكراري المتجمع الصاعد أو المنحنى التكراري المتجمع النازل على معلم متعامد، ثم رسم مستقيم موازي للمحور الأفقي انطلاقاً من النقطة التي

تمثل ترتيب الوسيط المعروف بالمعادلة 4-39 أعلاه إنطلاقاً من المحور العمودي، وتكون نقطة تقاطع هذا المستقيم مع المنحنى المتجمع، هي التي تحدد قيمة الوسيط، وذلك بإنزال شاقول من هذه النقطة على المحور الأفقي، فنجد قيمة الوسيط عند نقطة تقاطعهما.

مثال 4-17: أوجد وسيط بيانات المثال 4-15 أعلاه باستخدام الطريقة البيانية.

الإجابة: بتطبيق مبدأ إيجاد الوسيط بالطريقة البيانية أعلاه وإعتقاداً على التكرار المتجمع الصاعد كما هو معروض في الجدول 4-16، نجد من خلال الشكل 4-1 أن $Mé$ يساوي تقريباً نفس قيمة الوسيط كما حسبت بالطريقتين السابقتين وهي:

$$Mé = 23.33$$



شكل 4-1

يمكن إيجاد الوسيط بيانياً أيضاً، بإنزال شاقول من نقطة تقاطع المنحنيين المتجمعين الصاعد والنازل، على المحور الأفقي، إذ أن نقطة تقاطع هذا الشاقول مع المحور الأفقي هي القيمة الوسيطة، وذلك ما يوضحه المثال التالي:

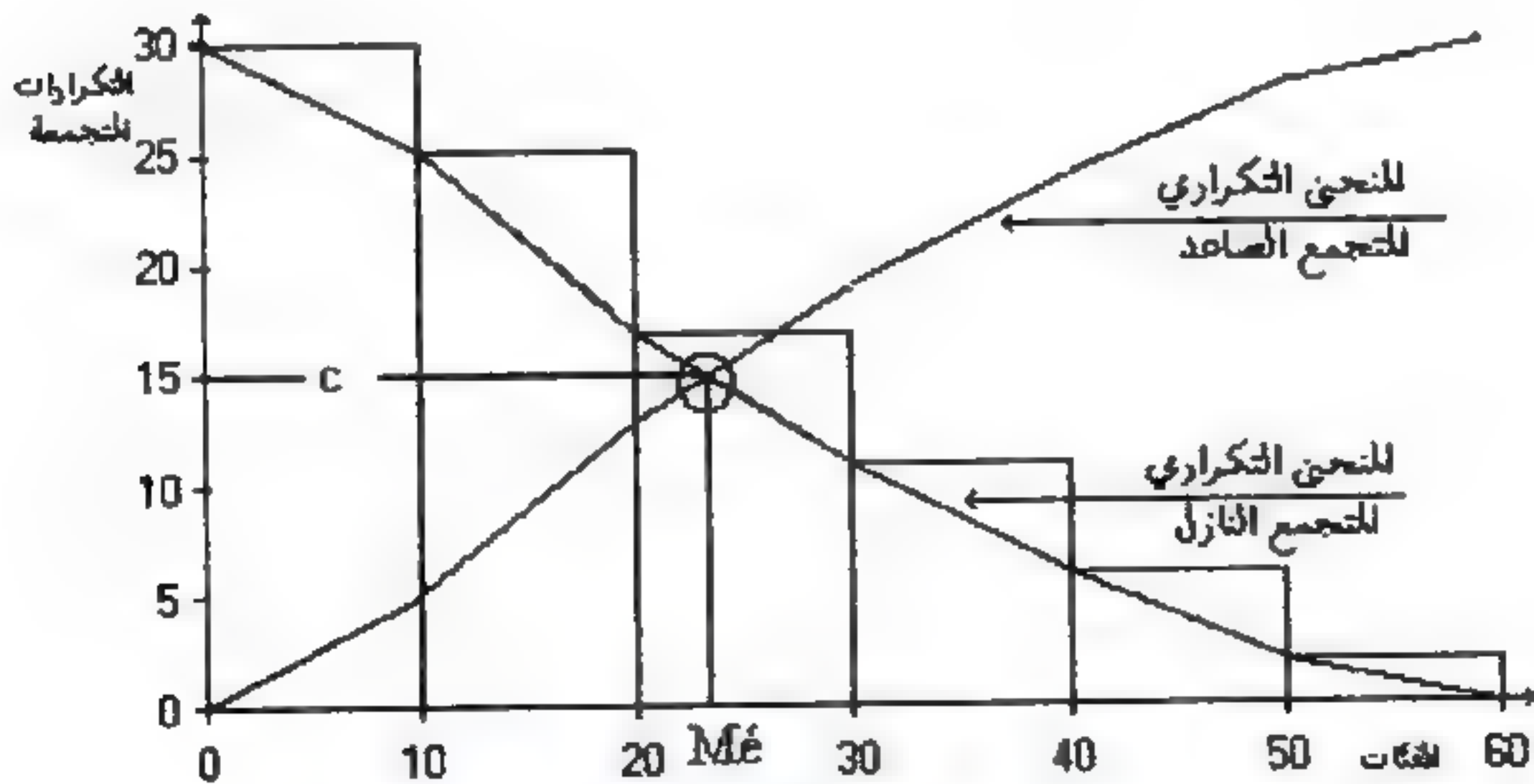
مثال 4-18: أوجد وسيط يانان المثال 4-15، من خلال نقطة التقاطع بين المنحنيين المتجمعين الصاعد والنازل.

الإجابة: التكرار المتجمع الصاعد معروض في الجدول رقم 4-16 أعلاه، أما التكرار المتجمع النازل فهو معروض في الجدول 4-17 أدناه.

ت.م.ن	الحد الأدنى	f_i	الفئات	i
30	المجموع	5	-10	1
25	10 فأكثر	8	10-20	2
17	20 فأكثر	6	20-30	3
11	30 فأكثر	5	30-40	4
6	40 فأكثر	4	40-50	5
2	50 فأكثر	2	50-60	6
		30		مج

جدول 4-17

برسم المنحنيين المتجمعين الصاعد والنازل على نفس المعلم، نجدهما متقاطعين بالفعل، عند نقطة شاقولها يتقاطع مع المحور الأفقي عند نقطة تمثل قيمة الوسيط $Mé = 23.33$ ، وذلك ما يوضحه الشكل التالي:



شكل 4-2

4- خصائص الوسيط : على عكس الوسط الحسابي، فإن الوسيط لا يتأثر بالقيم الشاذة والمتطرفة، لأنه لا يعتمد على القيم في حد ذاتها، بل على ترتيبها، وهو بالتالي أقل أهمية من الوسط الحسابي لكونه لا يحقق إلا بعضاً من شروط يول، و عموماً يغلب استخدامه في البيانات التي لا تعرف قيمها، ولكن ترتيبها معروف، كما يستخدم أيضاً في البيانات الناقصة وكذا في التوزيعات التكرارية المفتوحة.

ثالثاً: المنوال

تعريفه 4-4: يعرف المنوال لمجموعة من البيانات بأنه الفئة الأكثر تكراراً بين مجموعة القيم، ويتم حسابه كما يلي :

1- البيانات غير المبوبة والبيانات المبوبة التي طول فئاتها محدود : منوال مثل هذه البيانات هو الفئة الأكثر تكراراً.

2- البيانات المبوبة التي مدي فئاتها أكبر من الصفر : لإيجاد المنوال يتم استخدام إحدى الطرق التالية :

أ- الطريقة الأولى : وهي أبسط الطرق، وتسمى بطريقة مركز الفئة المنوالية، وفيها تكون قيمة المنوال هي الوسط الحسابي للفئة المنوالية.

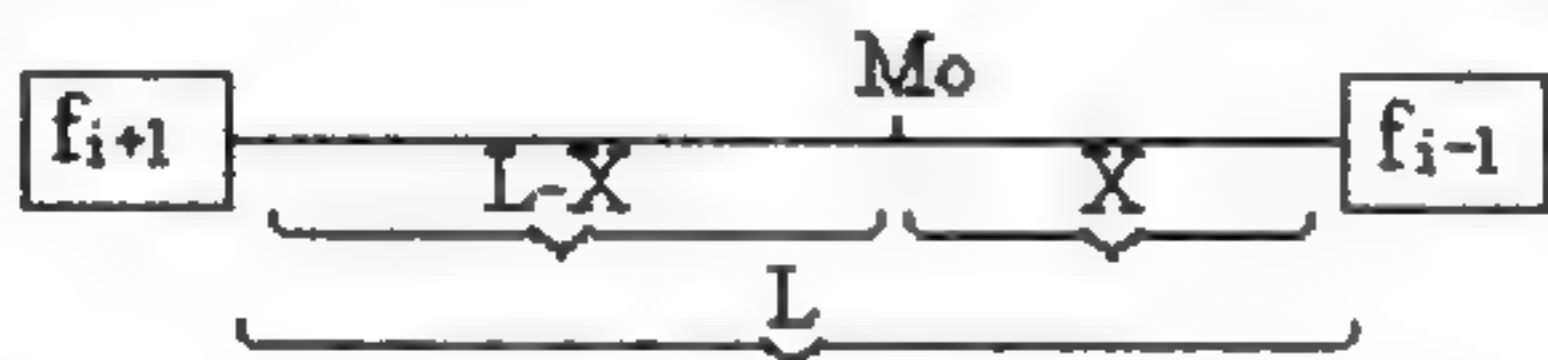
ب- الطريقة الثانية : تسمى بطريقة الرافعة، وفيها يتم إيجاد المنوال باستخدام المعادلة التالية :

$$Mo = d + \frac{f_{i+1}}{f_{i+1} + f_{i-1}} \times L \quad 42-4$$

حيث :

Mo	الموال
d	الحد الأدنى للفئة المتوالية
L	طول الفئة المتوالية
f_{i+1}	التكرار اللاحق لتكرار الفئة المتوالية
f_{i-1}	التكرار السابق لتكرار الفئة المتوالية

تسمى هذه الطريقة بطريقة الرافعة لإعتمادها على مبدأ الرافعة المعروفة في الفيزياء، و مبدأها أن النقطة التي يحدث عندها توازن الرافعة هي التي يتحقق عندها تساوي الكتلة اليمنى في الذراع الأيمن مع الكتلة اليسرى في الذراع الأيسر، فإذا افترضنا أن الموال هو النقطة التي يحدث عندها التوازن داخل الفئة المتوالية تتجاذبه كتلتان هما التكرار السابق لتكرار الفئة المتوالية و التكرار اللاحق للفئة المتوالية، فإنه يجب أن يساوي إلى الحد الأدنى للفئة المتوالية مضافا إليه مقدار ثابت بين الحد الأدنى ونقطة التوازن، ويمكن تصور ذلك من خلال الشكل التالي:



من الشكل أعلاه ينبغي البحث عن القيمة X التي يجب إضافتها إلى الحد الأدنى للفئة المتوالية d للحصول على القيمة المتوالية. بتطبيق مبدأ الرافعة وإعتمادا على الشكل السابق نجد:

$$f_{i-1} \times X = f_{i+1} (L - X)$$

بفك المعادلة نجد أن القيمة X هي :

$$X = \frac{f_{i+1}}{f_{i+1} + f_{i-1}} \times L$$

و بالتالي تكون القيمة المتوالية حسب المعادلة 4-42 أي:

$$Mo = d + \frac{f_{i+1}}{f_{i+1} + f_{i-1}} \times L$$

مثال 4-19: أوجد منوال البيانات التالية :

i	1	2	3	4	5
الفئات	3-5	5-7	7-9	9-11	11-13
f_i	3	10	20	15	7

جدول 4-18

من الجدول نجد أن الفئة المنوالية، أي الفئة الأكثر تكرارا هي : (7-9) وبالتالي نجد :

$$f_{i-1} = 10 \quad f_{i+1} = 15 \quad L = 2 \quad d = 7$$

بتطبيق المعادلة رقم : 4-42 نجد :

$$Mo = 7 + \frac{15}{15 + 10} \times 2 = 8.2$$

$$Mo = 8.2$$

أي أن القيمة المنوالية هي :

ج- الطريقة الثالثة : وتسمى بطريقة الفروقات أو طريقة بيرسون، وفيها يعطى المنوال بالقاعدة التالية :

$$Mo = d + \frac{d_{i-1}}{d_{i+1} + d_{i-1}} \times L \quad 4-43$$

حيث :

d	الحد الأدنى للفئة المنوالية
L	طول الفئة المنوالية
d_{i-1}	الفرق بين تكرار الفئة المنوالية و تكرار الفئة السابقة لها
d_{i+1}	الفرق بين تكرار الفئة المنوالية و تكرار الفئة اللاحقة لها

مثال 4-20: أوجد منوال بيانات المثال 4-19 باستخدام طريقة الفروقات.

الإجابة: بتطبيق المعادلة 4-43 نجد :

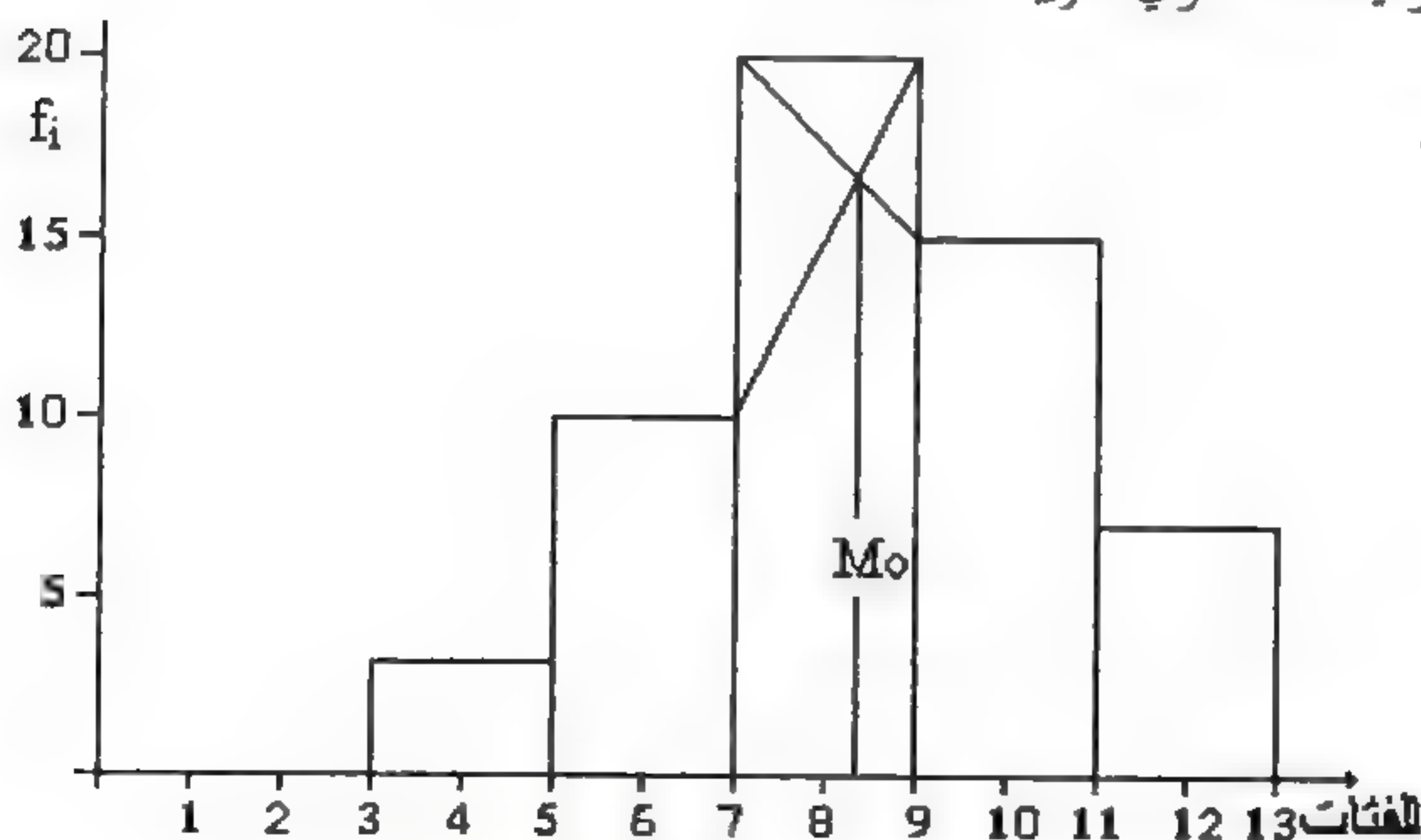
$$d_{i+1}=20-15=5 \quad d_{i-1} = 20-10=10 \quad L=2 \quad d=7$$

$$Mo = 7 + \frac{10}{10+5} \times 2 = 8.33$$

ومنه يكون :

يلاحظ أن هذه الطريقة لاتعطي قيم متطابقة تماما مع القيم التي تعطيها الطريقة السابقة، لكنها متقاربة في أغلب الأحيان.

د-الطريقة الرابعة: وهي الطريقة البيانية، وفيها يتم رسم المدرج التكراري ثم يتم أولا الوصل بين النقطة التي تمثل الحد الأعلى للفئة السابقة للفئة المتواليية والنقطة التي تمثل الحد الأعلى للفئة المتواليية، ثم ثانيا بين النقطة التي تمثل الحد الأدنى للفئة المتواليية والنقطة التي تمثل الحد الأدنى للفئة اللاحقة للفئة المتواليية، يتقاطع الخطان في نقطة شاقولها على المحور الأفقي يمثل القيمة المتواليية، وذلك كما هو واضح في الشكل أدناه، والذي هو تجسيد لبيانات المثال 4-19، حيث نجد القيمة المتواليية تساوي تقريبا 8.3.



شكل 3-4

3-خصائص المنوال : من خصائص المنوال، أنه أبسط مقاييس التزعة المركزية، وأنه لا يتأثر بالقيم المتطرفة، لأنه لا يأخذ

في الحسابان جميع القيم، كما يمكن إيجاد جميع التوزيعات بما فيها التوزيعات التكرارية المفتوحة، ومن مميزاته أيضا أنه لا يحتاج الى حسابات معقدة، الا في البيانات التكرارية التي مدى فئاتها أكبر من الصفر، وهو على عكس مقاييس التزعة المركزية الأخرى، إذ يمكن أن يكون لمجموعة من البيانات أكثر من منوال واحد، إضافة الى هذا فإنه المقياس الوحيد الذي يمكن تطبيقه على البيانات ذات الصفات النوعية.

4-العلاقة بين المنوال الوسيط الحسابي والوسيط : إذا كان لمجموعة من البيانات منوال واحد، ومنحنائها التكراري يقترب من التماثل، فإن قيمة الوسيط عموما تكون بين الوسيط الحسابي والمنوال، وتحقق المعادلة التالية بصفة تقريبية.

$$\frac{\bar{x} - Mo}{3} = \bar{x} - Me \quad 44-4$$

رابعا:الوسط الهندسي

تعريف 4-5: الوسط الهندسي لأية مجموعة من القيم، هو الجذر النوني لجداءات تلك القيم، وتختلف طريقة حسابه حسب طبيعة تقديم البيانات :

1-البيانات غير المبوبة : اذا كانت لدينا القيم $x_1, x_2, x_3, \dots, x_n$ ، فإن وسطها الهندسي هو :

$$G = \sqrt[n]{x_1 \times x_2 \times x_3 \times \dots \times x_n} \quad 45-4$$

واختصارا يكتب على شكل المعادلة 46-4 أدناه :

$$G = \left[\prod_{i=1}^n x_i \right]^{\frac{1}{n}} \quad 46-4$$

غير أنه لتسهيل الحسابات، ينبغي إدخال اللوغاريتم على الطرفين، وإجراء الحسابات اللازمة ثم تحويل النتيجة بعد ذلك إلى قيمتها الحقيقية، ويتم ذلك كما يلي :

المعادلة 4-45 تكتب كما يلي :

$$G = (x_1 \times x_2 \times x_3 \times \dots \times x_n)^{\frac{1}{n}}$$

بإدخال اللوغاريتم على الطرفين نجد :

$$\text{Log}G = \frac{1}{n} \text{Log}(x_1 \times x_2 \times x_3 \times \dots \times x_n)$$

حسب خواص اللوغاريتمات فإنه يمكن أن نكتب هذه العبارة كما يلي:

$$\text{Log}G = \frac{1}{n} (\text{Log}x_1 + \text{Log}x_2 + \text{Log}x_3 + \dots + \text{Log}x_n)$$

و منه نجد :

$$\text{Log}G = \frac{1}{n} \sum_{i=1}^n \text{Log} x_i \quad 47-4$$

و بالعودة إلى الأصل نجد :

$$G = 10^{\frac{1}{n} \sum_{i=1}^n \text{Log} x_i} \quad 48-4$$

حيث القيمة 10 هي أساس اللوغاريتم العشري، و يتم إستبدالها بالقيمة $e = 2.718$ في حالة إستخدام اللوغاريتم النيبيري.

مثال 4-21: أوجد الوسط الهندسي للبيانات التالية :

2 5 3 7 10 20

لإيجاد الوسط الهندسي نطبق المعادلة 4-45 أو 4-46، على النحو التالي:

$$G = \sqrt[6]{2 \times 5 \times 3 \times 7 \times 10 \times 20}$$

بإدخال اللوغاريتم على الطرفين نجد :

$$\text{Log}G = \frac{1}{6} (\text{Log}2 + \text{Log}5 + \text{Log}3 + \text{Log}7 + \text{Log}10 + \text{Log}20) = 0.77$$

و منه نجد:

$$G = 10^{0.77} = 5.89$$

2-البيانات المبوبة : في هذه الحالة يجب أخذ التكرارات بعين الاعتبار، لذلك فالوسط الهندسي يعطى بالمعادلة التالية :

$$G = \sqrt[n]{x_1^{f_1} \times x_2^{f_2} \times x_3^{f_3} \times \dots \times x_n^{f_n}} \quad 49-4$$

لإيجاد قيمة الوسط الهندسي يتم أيضا إدخال اللوغاريتم على طرفي المعادلة، أي إيجاد قيمة لوغاريتم الوسط الهندسي أولا، ثم إيجاد قيمة الوسط الهندسي بعد ذلك برفع القيمة 10 الى أس اللوغاريتم و ذلك كما يلي:

$$G = \left[x_1^{f_1} \times x_2^{f_2} \times x_3^{f_3} \times \dots \times x_n^{f_n} \right]^{\frac{1}{\sum f_i}}$$

وبادخال اللوغاريتم نجد :

$$\text{Log}G = \frac{1}{\sum f_i} \sum_{i=1}^n f_i \text{Log}x_i$$

و منه يكون :

$$G = 10^{\frac{1}{\sum f_i} \sum_{i=1}^n f_i \text{Log}x_i} \quad 50-4$$

هذا فيما لو استخدم اللوغاريتم العشري، أما عندما نستخدم اللوغاريتم النيبيري، فيتم إستبدال الرقم 10 بـ : $e = 2.718$.
مثال: 4-22: أوجد الوسط الهندسي للبيانات التالية :

i	x_i	f_i
1	5	20
2	8	25
3	10	15
4	15	10
مج		70

جدول 4-19

باستخدام المعادلة 4-50 و بمساعدة الجدول التالي :

i	x_i	f_i	$f_i \cdot \log x_i$
1	5	20	13.98
2	8	25	22.58
3	10	15	15.00
4	15	10	11.76
مج		70	63.32

جدول 4-20

$$\log G = \frac{1}{70} \times 63.32 = 0.9045 \quad \text{نجد :}$$

$$G = 10^{0.9045} = 8.03 \quad \text{أي :}$$

ويتم إيجاد الوسط الهندسي للبيانات التي مدى فئاتها أكبر من الصفر، بنفس الطريقة مع إستبدال القيم x_i بمراكز الفئات وذلك في المعادلة 4-50 أو 4-49.

ويستخدم الوسط الهندسي في البيانات التي تشكل أو تكاد تشكل متوالية هندسية، كبيانات تطور عدد السكان، أو بيانات المبالغ المستثمرة وفق فائدة مركبة، أما في غير ذلك فهو نادر الاستخدام لصعوبة حسابه.

ومن خواصه أن حاصل ضرب مجموعة من القيم لا يتغير إذا استبدلت كل قيمة من هذه القيم بالوسط الهندسي، فإذا كانت لدينا القيم : 2، 5، 15 مثلاً فإن جداولها هو : $2 \times 5 \times 15 = 150$ بينما وسطها الهندسي هو :

$$G = \sqrt[3]{2 \times 5 \times 15} = 5.31$$

عند إستبدال كل قيمة بالوسط الهندسي نجد :

$$150 = 5.31 \times 5.31 \times 5.31$$

و يكون الوسط الهندسي دائماً أقل من الوسط الحسابي، ولا يتساوى معه إلا إذا كانت جميع قيم الظاهرة متساوية.

كما أن الوسط الهندسي يأخذ بعين الاعتبار جميع مفردات القيم، و لا يتأثر بالقيم الشاذة، و هو يحقق بعضا من شروط يول، ومن عيوبه أنه لا يمكن حسابه في حالة بيانات التوزيعات التكرارية المفتوحة أو التي يكون جدها سالبا.

خامسا: الوسط التربيعي

تعريفه 4-6 : الوسط التربيعي لأية مجموعة من القيم، هو الجذر التربيعي للوسط الحسابي لمربعات تلك القيم، ويتم حسابه حسب طبيعة البيانات كما يلي:

1- البيانات غير المبوبة : اذا كانت لدينا البيانات :

$X_1, X_2, X_3, \dots, X_n$

فإن وسطها التربيعي يعطى كما يلي :

$$Q = \sqrt{\frac{X_1^2 + X_2^2 + X_3^2 + \dots + X_n^2}{N}} \quad 51-4$$

واختصارا يكتب :

$$Q = \sqrt{\frac{\sum_{i=1}^n X_i^2}{N}} \quad 52-4$$

مثال 4-23: أوجد الوسط التربيعي لبيانات المثال 4-21.

لايجاد الوسط التربيعي نبحث عن مربعات القيم ثم نوجد الجذر التربيعي لوسطها الحسابي كما يلي :

i	1	2	3	4	5	6	مجموع
X_i	2	5	3	7	10	20	47
X_i^2	4	25	9	49	100	400	587

جدول 4-21

$$Q = \sqrt{\frac{587}{6}} = 9.89 \quad \text{ومنه يكون :}$$

2- البيانات المربوبة : في هذه الحالة يجب أخذ التكرارات بعين الاعتبار، بحيث يكون الوسط التربيعي على النحو التالي :

$$Q = \sqrt{\frac{\sum_{i=1}^n x_i^2 f_i}{\sum_{i=1}^n f_i}} \quad 53-4$$

مثال 4-24: أوجد الوسط التربيعي لبيانات المثال 4-22. بتطبيق المعادلة رقم: 53-4 وبمساعدة الجدول التالي :

i	x_i	f_i	x_i^2	$x_i^2 \cdot f_i$
1	5	20	25	500
2	8	25	64	1600
3	10	15	100	1500
4	15	10	225	2250
مج		70		5850

جدول 4-22

$$Q = \sqrt{\frac{5850}{70}} = 9.14 \quad \text{نجد :}$$

وفي حالة البيانات المربوبة التي مدى فئاتها أكبر من الصفر يتم كذلك إستبدال القيم x_i بمراكز الفئات في المعادلة 53-4. من خواص الوسط التربيعي أنه دائماً أكبر من الوسط الحسابي، إلا إذا كانت جميع القيم متساوية، وهو يحقق جزئياً بعضاً من شروط يول، كما أنه قليل الاستخدام ويمكن الإستعانة به في حساب الانحراف المعياري كما سيأتي في الفصل الموالي.

ماددا : الوسط التوافقي

تعريف 4-7: الوسط التوافقي لمجموعة من القيم، هو مقلوب الوسط الحسابي لمقلوب تلك القيم، ويتم إيجاده حسب طبيعة البيانات كما يلي:

1- البيانات غير المبوبة: إذا كانت لدينا البيانات: $x_1, x_2, x_3, \dots, x_n$ ، فإن وسطها التوافقي يعطى كما يلي :

$$H = \frac{N}{\sum_{i=1}^n \left[\frac{1}{x_i} \right]} \quad 54-4$$

مثال 4-25: أوجد الوسط التوافقي لبيانات المثال 4-21: بتطبيق المعادلة رقم 54-4 وبمساعدة الجدول التالي:

i	1	2	3	4	5	6	مج
x_i	2	5	3	7	10	20	47
خطأ! الإشارة المرجعية غير معروفة.	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{1}{7}$	$\frac{1}{10}$	$\frac{1}{20}$	1.326

جدول 4-23

$$H = \frac{6}{1.326} = 4.52 \quad \text{نجد :}$$

2- البيانات المبوبة: إذا كانت لدينا البيانات: $x_1, x_2, x_3, \dots, x_n$ تكراراتها على التوالي: $f_1, f_2, f_3, \dots, f_n$ فإن وسطها التوافقي يعطى كما يلي:

$$H = \frac{\sum_{i=1}^n f_i}{\sum_{i=1}^n \left[\frac{1}{x_i} \right] f_i} \quad 55-4$$

مثال 4-26: أوجد الوسط التوافقي لبيانات المثال 4-22. باستخدام المعادلة رقم 55-4 وبمساعدة الجدول التالي :

i	x_i	f_i	$\frac{1}{x_i}$	$\frac{1}{x_i} f_i$
1	5	20	0.200	4.000
2	8	25	0.125	3.125
3	10	15	0.100	1.500
4	15	10	0.067	0.670
مج		70		9.295

جدول 4-24

$$H = \frac{70}{9.295} = 7.53 \quad \text{نجد :}$$

نتيجة لصعوبة حساب الوسط التوافقي، فإنه أقل مقاييس التزعة المركزية إستخداما، و يقتصر إستخدامه أحيانا على إيجاد متوسطات الأسعار.

3- العلاقة بين الوسط التوافقي والوسط الحسابي، الوسط الهندسي: إذا كانت لدينا مجموعة من القيم وحسبت أوساطها الحسابي، الهندسي والتوافقي، فإننا سوف نجد العلاقة التالية بينها :

$$H \leq G \leq \bar{x} \quad 56-4$$

مايجدر الإشارة إليه أخيرا هو أن جميع مقاييس التزعة المركزية تأخذ نفس وحدات قياس المعلومات الأولية، سواء كانت وحدات القياس هذه طبيعية بسيطة أو مركبة أونقدية، كما نشير أيضا الى أن الوسطين، الحسابي والتربيعي يتأثران بالقيم الكبيرة، اذ أنهما أكبر تحيزا إليها، بينما الوسطين التوافقي والهندسي أكثر تحيزا للقيم الصغيرة.

هذه هي مقاييس التزعة المركزية الأكثر شهرة، وهناك مقاييس جزئية أخرى، تهتم بترتيب القيم، لذلك نصلح على تسميتها بمقاييس الترتيب أو مقاييس الوضع ومنها الربيعات، العشيرات والمئيات.

سابعا: الربيعيات : كل مجموعة من البيانات يمكن تقسيمها الى أربعة أقسام متساوية بعد ترتيبها تصاعديا، يفصل بين كل قسم، ما

يسمى بالربيع، ويتعارف على ربيعين أساسيين، هما الربيع الأدنى، ويسمى أيضا الربيع الأول، والربيع الأعلى، ويسمى كذلك الربيع الثالث، أما الربيع الأوسط فهو عبارة عن الوسيط كما عرفناه سابقا، ونرمز للربيعيات بـ: Q_i ، حيث: $i=1, 2, 3$.

1- الربيع الأدنى:

تعريف 4-8 : الربيع الأدنى لمجموعة من القيم، هو القيمة التي يكون قبلها 25 % على الأكثر وبعدها 75 % على الأكثر، من إجمالي عدد القيم، بعد ترتيبها تصاعديا، ويحسب ترتيبه كما يلي:

أ- حالة البيانات غير المبوبة:

$$c_1 = N \cdot \frac{25}{100} \quad 57-4$$

حيث N : عدد القيم.

ب- حالة البيانات المبوبة:

$$c_1 = \sum_{i=1}^n f_i \frac{25}{100} \quad 58-4$$

2- الربيع الأعلى:

تعريف 4-9 : الربيع الأعلى لمجموعة من البيانات، هو القيمة التي يكون قبلها 75 % على الأكثر وبعدها 25 % على الأكثر من إجمالي عدد القيم، بعد ترتيبها تصاعديا، ويحسب ترتيبه كما يلي :

أ- حالة البيانات غير المبوبة :

$$c_3 = N \cdot \frac{75}{100} \quad 59-4$$

ب- حالة البيانات المبوبة :

$$c_3 = \sum_{i=1}^n f_i \cdot \frac{75}{100} \quad 60-4$$

مثال 4-27: أوجد الربيع الأدنى و الربيع الأعلى للبيانات التالية:

2 4 5 8 10 13 15

الإجابة : بتطبيق المعادلتين 4-58 و 4-59 نجد ترتيب

الربعين على التوالي : $c_1 = 1.75$ $c_3 = 5.25$

لذلك يجب أن يكون قبل الربع الأول على الأكثر قيمة واحدة و بعده 5 قيم على الأكثر، وبالتالي يكون :

$$Q_1 = 4$$

كما يجب أن يكون قبل الربع الأعلى 5 قيم على الأكثر و بعده قيمة واحدة على الأكثر، وبالتالي يكون:

$$Q_3 = 13$$

ملاحظة: بالنسبة للبيانات المبوبة أنظر المثال 4-28 في نهاية هذا الفصل.

ثامنا، العشرية : كل مجموعة من البيانات يمكن تقسيمها

الى عشرة أقسام متساوية، بعد ترتيبها تصاعديا، يفصل بين كل قسم ما يسمى بالعشيرة، وهناك ما يسمى بالعشير الأول، الثاني، الثالث... و التاسع، ونرمز للعشيريات بالرمز : D_i ، حيث :

$$i=1,2,3,\dots,9$$

تعريفه 4-10: عشير مجموعة من البيانات هو القيمة التي تفصل بين أقسام هذه المجموعة بعد تجزئتها الى عشرة أجزاء متساوية، وعلى هذا فالعشير الأول هو القيمة التي يكون قبلها عشر البيانات وبعدها تسعة أعشارها، والعشير الثاني هو القيمة التي يكون قبلها عشري البيانات وبعدها ثمانية أعشارها... ويحسب ترتيب كل عشير كما يلي :

$$c_i = N \cdot \frac{i}{10} \quad 61-4$$

حيث : i : رقم العشيرة، N : عدد القيم. هذا في حالة البيانات غير المبوبة، أما في حالة البيانات المبوبة، فيتم إستبدال N بمجموع التكرارات، ويكون :

$$c_i = \sum_{j=1}^n f_j \cdot \frac{i}{10} \quad 62-4$$

ثامنا، المئوية : يمكن كذلك تقسيم كل مجموعة من البيانات الى مائة قسم متساو بعد ترتيبها تصاعديا، يفصل بين كل قسم ما يسمى

بالمؤين، وهناك ما يسمى بالمؤين الأول، الثاني، الثالث... و التاسع والتسعون، ونرمز للمؤين بالرمز : C_i ، حيث : $i=1, 2, 3, \dots, 99$

تعريف 4-11: مؤين مجموعة من البيانات هو القيمة التي تفصل بين أقسام هذه المجموعة بعد تجزئتها الى مائة جزء متساو، وعلى هذا فالمؤين الأول هو القيمة التي يكون قبلها واحد على مائة من البيانات وبعدها تسعة وتسعون على مائة من تلك البيانات، والمؤين العاشر هو القيمة التي يكون قبلها عشرة على مائة من البيانات وبعدها تسعون على مائة من تلك البيانات... و يحسب ترتيب كل مؤين كما يلي :

$$c_i = N \cdot \frac{i}{100} \quad 63-4$$

حيث : i : رقم المؤين، N : عدد القيم. هذا في حالة البيانات غير المبوبة، أما في حالة البيانات المبوبة، فيتم إستبدال N بمجموع التكرارات، ويكون :

$$c_i = \sum_{i=1}^n f_i \cdot \frac{i}{100} \quad 64-4$$

و ما تجدر الإشارة اليه هو أن، طريقة إيجاد كل من الربيعيات والعشيرات والمؤينات، تتم بنفس طرق إيجاد الوسيط، مع إستبدال ترتيب الوسيط c ، بترتيب الربيعيات والعشيرات والمؤينات، حسب الحالة.

وعند إستخدام الطريقة البيانية، لإيجاد، كل من الربيعيات والعشيرات والمؤينات، فإنه يتم رسم المنحنى التكراري المتجمع النازل أو الصاعد على معلم متعامد، وإنطلاقاً من نقطة ترتيب إما الربيع أوالعشير أو المؤين، يتم إمداد خط مستقيم موازي للمحور الأفقي، وتكون نقطة تقاطع هذا المستقيم مع المنحنى المتجمع النازل أو الصاعد هي التي تحدد الربيع أو العشير أوالمؤين حسب الحالة، وذلك بإنزال شاقول من نقطة

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

التقاطع تلك على المحور الأفقي، فنحصل بذلك على، الربيع أو العشير أو المؤين.

ملاحظة: العلاقة بين الوسيط الربيعيات، العشيريات والمؤينات: إذا حسبت ربيعيات وعشيريات ومؤينات مجموعة من البيانات، فإننا نجد العلاقة التالية بينها:

$$M_e = Q_2 = D_5 = C_{50}$$

$$Q_1 = C_{25}$$

$$Q_3 = C_{75}$$

مثال 4-28: أوجد الربيع الأول والثالث، والعشير الخامس، والمؤين العاشر والخمسون و الوسيط، للبيانات التالية :

i	الفئات	f _i
1	5-10	3
2	10-15	5
3	15-20	10
4	20-25	6
5	25-30	4
6	30-35	4
مج		32

جدول 4-25

وذلك باستخدام :

- 1- الطريقة الأولى لإيجاد الوسيط في حالة البيانات المبوبة التي طول فئاتها أكبر من الصفر (معادلة رقم 4-40).
- 2- الطريقة البيانية.

الإجابة : 1- المعادلة 4-40 المطلوب إستخدامها هي :

$$M_e = d + \frac{c - f_{i-1}^+}{f_{i+1}^+ - f_{i-1}^+} \times L$$

غير أنه ينبغي إستبدال ترتيب الوسيط، بترتيب المقياس المرغوب في حسابه، ولأجل ذلك نوجد التكرار المتجمع الصاعد، وهو كما يلي:

i	الفئات	f _i	الحدود الدنيا	
1	10-5	3	أقل من 10	3
2	15-10	5	أقل من 15	8
3	20-15	10	أقل من 20	18
4	25-20	6	أقل من 25	24
5	30-25	4	أقل من 30	28
6	35-30	4	أقل من 35	32
مج		32		

جدول 4-26

1- الربيع الأول :

ترتيب الربيع الأول : بتطبيق المعادلة رقم 4-58 نجد: $c_1 = 8$
ومنه تكون قيمة الربيع الأول هي :

$$Q_1 = 15 + \frac{8-8}{18-8} \times 5 = 15$$

ب- الربيع الثالث :

ترتيب الربيع الثالث : بتطبيق المعادلة 4-60 نجد: $c_3 = 24$
ومنه تكون قيمة الربيع الثالث هي :

$$Q_3 = 25 + \frac{24-24}{28-24} \times 5 = 25$$

ج- العشير الخامس :

ترتيب العشير الخامس : بتطبيق المعادلة 4-62 نجد :
 $c_5 = 16$ ومنه تكون قيمة العشير الخامس هي :

$$D_5 = 15 + \frac{16-8}{18-8} \times 5 = 19$$

د- المؤين العاشر :

ترتيب المؤين العاشر : بتطبيق المعادلة 4-64 نجد: $c_{10} = 3.2$ ومنه تكون
قيمة المؤين العاشر هي :

$$C_{10} = 10 + \frac{3.2 - 3}{8 - 3} \times 5 = 10.2$$

هـ- المؤيّن الخمسون :

ترتيب المؤيّن الخمسون : بتطبيق المعادلة 4-64 نجد:

$C_{50} = 16$ ، و منه تكون قيمة المؤيّن الخمسون هي :

$$C_{50} = 15 + \frac{16 - 8}{18 - 8} \times 5 = 19$$

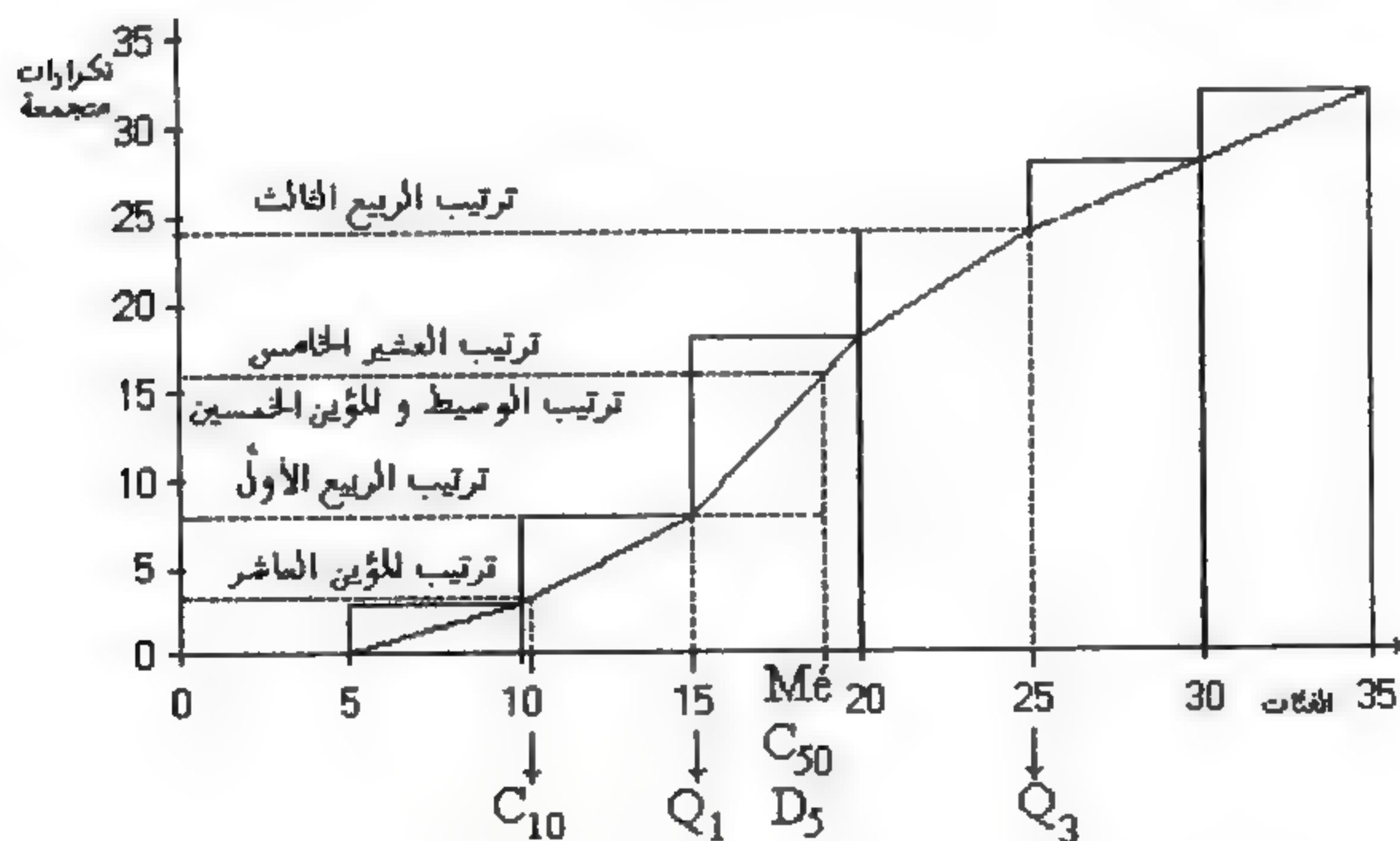
و- الوسيط :

ترتيب الوسيط هو : $c = 16$

ومنه تكون قيمة الوسيط هي :

$$Me = 15 + \frac{16 - 8}{18 - 8} \times 5 = 19$$

2- بتطبيق فكرة إيجاد الربيعيات والعشيرات والمئينات بيانياً ومن خلال الشكل 4-3 أدناه، نحصل تقريباً على نفس القيم المحصل عليها أعلاه.



شكل 4-3

تمارين

تمرين 1: البيانات التالية تظهر الأجور الأسبوعية لعمال مصنع ما بمئات الدينارات.

34	13	54	24	35	30	20	52
47	45	31	10	18	22	33	45
35	34	28	24	14	17	25	50
30	31	36	23	50	46	45	41

المطلوب:

- 1- أوجد الوسط الحسابي لهذه البيانات.
- 2- بوب البيانات أعلاه في جدول تكراري مستمر يجعل طول الفئة: $L = 500$
- 3- من الجدول المحصل عليه في المطلوب 2، أوجد الوسط الحسابي وقارنه بالوسط الحسابي المحصل عليه من المطلوب 1.
- 4- أوجد الوسيط قبل التبويب وبعده.
- 5- أوجد المنوال قبل التبويب وبعده.
- 6- أوجد الوسط الهندسي، الوسط التريبي قبل التبويب وبعده.
- 7- أوجد: - الربيعان الأول و الثالث - المؤي الخامس والعشرون - المؤي الخمسون - المؤي الخامس والسبعون - العشير الخامس وذلك بعد التبويب، باستخدام المعادلة 4-40، وباستخدام الطريقة البيانية.
- 8- قارن النتائج المحصل عليها من السؤال 7.

تمرين 2: الجدول التالي يظهر توزيع عمال مصنع ما حسب الأجور الأسبوعية، بمئات الدينارات.

الأجور	32-30	34-32	36-34	38-36	40-38	42-40	44-42
العمال	10	20	15	45	30	16	7

المطلوب:

- 1- أوجد متوسط الأعمار باستعمال طريقة الوسط الفرضي.
- 2- أوجد كل مقاييس التزعة المركزية الأخرى.
- 3- أثبت حسابيا تساوي كل من الوسيط، العشير الخامس، المؤين الخمسون.
- تحويل 3: البيانات التالية تظهر توزيع سكان الجزائر المقيمين، حسب فئات الأعمار و الجنس سنة 2000 بالآلاف.

المصدر: الجزائر بالأرقام العدد 31 الديوان الوطني للإحصائيات / www.ons.dz

الفترة/سنة	ذكور	إناث	الفترة/سنة	ذكور	إناث
0-4	1534	1495	45-49	626	611
5-9	1799	1720	50-54	441	447
10-14	1933	1860	55-59	346	362
15-19	1890	1817	60-64	316	334
20-24	1617	1571	65-69	267	283
25-29	1352	1331	70-74	185	194
30-34	1149	1136	75-79	112	115
35-39	930	918	80 وماكثر	112	125
40-44	751	744	مجموع	15360	15026

المطلوب:

- 1- قدم هذه البيانات في جدول تكراري مستمر يجعل طول الفئة: $L = 10$
- 2- من الجدول المحصل عليه قم بما يلي:
 - أ- أوجد الوسط الحساب لأعمار كل من الذكور والإناث.
 - ب- أوجد الوسيط و المنوال. ج- قارن المقاييس المحصل عليها. ماذا تستنتج.
 - د- أوجد وسيط الأعمار باستعمال الطريقة البيانية.
- 4- قارن النتائج وقدم الإستنتاج.

تمرين 4: البيانات التالية تظهر العمر المتوسط لأول زواج حسب الجنس والولاية في الجزائر، حسب مصادر الديوان الوطني للإحصائيات.

الولاية	ذكور	إناث	الولاية	ذكور	إناث	الولاية	ذكور	إناث	الولاية	ذكور	إناث
أدرار	26.2	20.1	فستنية	29.1	26.2	تلمسان	28.6	24	تندوف	27.9	20.8
الشلف	25.0	22.5	المدية	26.1	22.0	تيارت	26.6	22	تسميلت	26.9	21.1
الأغواط	27.3	22.3	مستعام	27.0	23.0	ت. وزور	27.6	23.3	الواد	25.8	20.6
أ. بوققي	27.4	24.4	المسيلة	25.7	21.1	الجزائر	30.7	27.2	خنشلة	27.4	23.9
بابة	27.0	23.5	معسكر	27.6	22.9	الجلقة	24.9	19.6	س. اهراس	27.6	23.9
بجاية	26.6	22.1	ورقلة	26.3	21.1	حيجل	27.7	24.1	تيازة	28.1	24.7
بسكرة	27.0	23	وهران	29.2	25.1	سطيف	26.4	22.6	ميلة	27.4	24.4
بشار	28.0	22.9	البيص	27.4	21.8	سعيدة	27.4	22.5	ع. الدعلي	26.8	22.8
البليدة	28.7	24.7	اليزي	27.5	20.4	سكيكدة	28.8	25.2	النعام	27.8	22.5
البويرة	26.7	22.5	ب. بو عريير	25.5	21.6	س. بلعاس	28.5	23.3	لموشنت	29.0	24.6
تمراست	27.3	20.4	بومرداس	29.2	25.0	عنابة	29.5	25.9	غرداية	25.8	21.2
تيسة	27.2	23	الطارف	28.0	24.6	قانة	28.6	25.5	غليزان	26.4	22.2

المطلوب:

- 1- أوجد المتوسط الوطني لسن أول زواج بالنسبة للجنسين.
- 2- أوجد متوسط عمر أول زواج للجنسين في المدن الكبرى: وهران، الجزائر، قسنطينة، عنابة.
- 3- أوجد متوسط عمر أول زواج للجنسين في كل من ولايات الشمال، ولايات الجنوب. ماذا تلاحظ؟.

تمرين 5: إذا كان : $\sum_{i=1}^6 x_i = -4$, $\sum_{i=1}^6 x_i^2 = 10$

إحسب نتائج العبارات التالية:

أ- $\sum_{i=1}^6 (2x_i + 5)$ ب- $\sum_{i=1}^6 x_i (x_i - 1)$ ج- $\sum_{i=1}^6 x_i \sum_{i=1}^6 (x_i - 5)^2$

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

تمرين 6: إذا كانت لدينا مجموعة من القيم وسطها معلوم، أثبت أنه:

1- إذا طرح مقدار ما C من كل قيمة من تلك القيم فإن الوسط الحسابي

للقيم الجديدة يكتب كما يلي : $\bar{x} = \bar{x} - C$

2- إذا قسمت كل قيمة من تلك القيم على العدد C ، فإن الوسط الحسابي

للقيم الجديدة يكتب كما يلي : $\bar{x} = \frac{\bar{x}}{C}$

حيث: \bar{x} : الوسط الحسابي للقيم الأصلية. \bar{x} : الوسط الحسابي للقيم الجديدة.

وذلك في حالتى البيانات المبوبة وغير المبوبة.

تمرين 7: إذا كانت لدينا مجموعتين من البيانات الأولى حجمها N_1 والثانية حجمها N_2 ، أثبت أنه إذا دجت المجموعتين فإن الوسط الحسابي للمجموعة الجديدة يكتب كما يلي:

$$\bar{x} = \frac{N_1 \cdot \bar{x}_1 + N_2 \cdot \bar{x}_2}{N_1 + N_2}$$

حيث \bar{x}_1 و \bar{x}_2 الوسطين الحسابيين للمجموعة الأولى و الثانية على التوالى.

الفصل الخامس

مقاييس التشتت

لتوضيح مفهوم التشتت نعطي المثال التالي :

مثال 5-1: قسمان دراسيان، كل قسم يحتوي على 5 طلبة، وكانت النتائج البيداغوجية لطلبة القسمين كما يلي :

القسم الأول	10	12	13	10	10
القسم الثاني	3	20	20	5	7

لو أخذنا الوسط الحسابي لنتائج القسمين لوجدناهما متساويين حيث:

$$\bar{x}_1 = \frac{55}{5} = 11$$

$$\bar{x}_2 = \frac{55}{5} = 11$$

يعني هذا أن المستوى البيداغوجي للقسمين متساو، غير أنه لو لاحظنا علامة كل طالب لوجدنا أن كل طلبة القسم الأول ناجحون، بينما لم ينجح في القسم الثاني سوى طالبين، فالحقيقة هي إذن أن مستوى القسمين غير متساو على الرغم من أن الوسط الحسابي لنتائجهما متساو، ولو لاحظنا الفرق بين أكبر علامة وأصغر علامة في القسم الأول لوجدناها: $(10-13) = 3$ ، بينما يقدر الفرق بين أكبر علامة وأصغر علامة في القسم الثاني بـ $(3-20) = 17$ ، أما أكبر إنحراف عن الوسط الحسابي في القسم الأول فهو: $(11-13) = 2$ ، وأكبر إنحراف عن الوسط الحسابي في القسم الثاني هو: $(11-20) = 9$ ويعني هذا أن نتائج القسم الثاني أكبر تشتتاً أي أكبر تباعداً عن بعضها وعن وسطها الحسابي، على عكس نتائج القسم الأول، فأنها أقل تشتتاً، والنتيجة هي أن الوسط الحسابي (وبقيّة مقاييس التّزعة

المركزية)، غير كافية وحدها لإعطاء خلاصة كافية عن البيانات، وذلك لأن الكثير منها يتأثر بالقيم المتطرفة، لذلك لابد من دراسة مدى تباعد القيم حتى يتمكن من إستخلاص نتائج أكثر واقعية، ومن هنا جاءت أهمية دراسة التشتت.

تعريف 1-5: التشتت هو مدى تباعد مجموعة القيم عن بعضها البعض أو عن القيمة التي تمثل مركز تلك المجموعة، ويقاس بعدة مقاييس هي :

أولاً: مدى التغير :

تعريف 2-5: يعرف مدى التغير أو المدى لأية مجموعة من القيم بأنه الفرق بين أعظم قيمة وأدنى قيمة من مجموعة القيم، سواء كانت هذه القيم مبوبة أو غير مبوبة، وهو معطى بالمعادلة 3-2 في الفصل الثاني، أي:

$$W = X_{Max} - X_{Min} \quad 1-5$$

مثال 2-5: من المثال 1-5 أوجد مدى التغير.

أكبر علامة في القسم الأول هي : 13 بينما أدنى علامة هي : 10 ومنه يكون مدى التغير للقسم الأول هو :

$$W_1 = 13 - 10 = 3$$

أكبر علامة في القسم الثاني هي : 20 بينما أدنى علامة هي : 3 ومنه يكون مدى التغير للقسم الثاني هو :

$$W_2 = 20 - 3 = 17$$

تعكس النتائج ما ذهبنا اليه في مقدمة هذا الفصل، وهو أن بيانات القسم الأول أقل تشتتاً وبالتالي فإن وسطها الحسابي يمثلها تمثيلاً مقبولاً، بينما بيانات القسم الثاني، متباعدة أي أكثر تشتتاً وبالتالي فإن وسطها الحسابي لا يمثلها تمثيلاً جيداً. غير أن مدى التغير يفقد أهميته نتيجة للعيوب التالية :

* أنه يتأثر بالقيم الشاذة لكونه لا يأخذ في الحسبان جميع القيم، فالبيانات التالية مثلاً: 3 ، 20 ، 25 ، 20 ، 23 ، 22 ، 90، جملتها يتراوح بين 20 و 25، غير أن مدى تغيرها كبير جداً وهو: $W=90-3=87$ ، فالمدى في مثل هذه الحالات لا معنى له، لأن تأثيره واضح بالقيمتين الشاذتين وهما: 3 و 90.

* مدى التغير لا يمكن استخدامه في حالة التوزيعات التكرارية المفتوحة، لأن حديها الأعلى و الأدنى يكونان غير معلومين.

ثانياً: الانحراف المتوسط :

تعريفه 3-5: الانحراف المتوسط هو الوسط الحسابي لفروقات القيم عن وسطها الحسابي بالقيمة المطلقة، وتختلف طريقة حسابه باختلاف طريقة تقديم البيانات :

1- للبيانات غير المبوبة : إذا كانت لدينا البيانات : $x_1, x_2, x_3, \dots, x_n$ فان انحرافها المتوسط يعطى بالمعادلة التالية :

$$e = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{N} \quad 2-5$$

مثال 3-5: البيانات التالية خاصة بالأجور الأسبوعية لعمال مؤسسة ما، بآلاف الدينارات، والمطلوب إيجاد انحرافها المتوسط.

50	70	80	50	60	50	70	50	70	90
70	100	50	80	60	70	50	60	60	80

الإجابة:

إيجاد الانحراف المتوسط يتطلب أولاً إيجاد الوسط الحسابي :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} = \frac{1320}{20} = 66 \text{ دج}$$

تطبيق المعادلة 2-5 يتطلب إيجاد الفروقات عن الوسط الحسابي كما هي واضحة أدناه :

24	4	-16	4	-16	-6	-16	14	4	-16
14	-6	-16	-16	4	-6	14	-16	34	4

ومعلوم أن مجموع هذه القيم معدوم، حسب أول خاصية من خصائص الوسط الحسابي، لذلك يتم أخذ القيم المطلقة لإيجاد الانحراف المتوسط، كما هو واضح في المعادلة 2-5، وبالتالي يكون :

$$e = \frac{240}{20} = 12.10^3$$

أي الانحراف المتوسط هو: 12.10^3 د ج.

2- **الانحراف المتوسط للبيانات المبوبة :** إذا كانت لدينا البيانات: $x_1, x_2, x_3, \dots, x_n$ تكراراتها على التوالي: $f_1, f_2, f_3, \dots, f_n$ فإن انحرافها المتوسط يعطى بالمعادلة التالية:

$$e = \frac{\sum_{i=1}^n |x_i - \bar{x}| f_i}{\sum_{i=1}^n f_i} \quad 3-5$$

مثال 4-5: بوب بيانات المثال 3-5 في جدول تكراري، ثم أحسب انحرافها المتوسط.

الإجابة: بتطبيق مبادئ التبويب كما هي واردة في الفصل الثاني نحصل على الجدول التالي :

i	x_i	f_i
1	50	6
2	60	4
3	70	5
4	80	3
5	90	1
6	100	1
مج		20

جدول 1-5

لإيجاد الانحراف المتوسط نحسب أولا الوسط الحسابي، ثم نطبق المعادلة 3-5 بمساعدة الجدول 2-5 أدناه:

i	x_i	f_i	$x_i f_i$	$ x_i - \bar{x} f_i$
1	50	6	300	96
2	60	4	240	24
3	70	5	350	20
4	80	3	240	42
5	90	1	90	24
6	100	1	100	34
مج		20	1320	240

جدول 2-5

$$\bar{x} = \frac{1320}{20} = 66.10^3 \quad \text{الوسط الحسابي هو:}$$

و يكون الانحراف المتوسط كما يلي :

$$e = \frac{240}{20} = 12.10^3 \text{ دج}$$

تجدر الإشارة الى أنه في حالة البيانات التي مدى فئاتها أكبر من الصفر، يتم استخدام مراكز الفئات c_i ، وتكون معادلة الانحراف المتوسط كما يلي :

$$e = \frac{\sum_{i=1}^n |c_i - \bar{x}| f_i}{\sum_{i=1}^n f_i} \quad 4-5$$

إن الخاصية الإيجابية للانحراف المتوسط هي أنه يأخذ في الحسبان جميع القيم، لذلك فدرجة تأثره بالقيم الشاذة ضعيفة، على عكس مدى التغير كما رأينا، وهو لا يحقق بعض شروط يول، ومن ذلك أنه لا يخضع للعمليات الجبرية، إذ أن مجموع انحرافات القيم عن وسطها الحسابي يكون معدوماً، لذلك تم استخدام القيمة المطلقة.

ما تجدر الإشارة إليه، هو أنه يمكن إيجاد الانحراف المتوسط بالنسبة لمختلف مقاييس التزعة المركزية الأخرى، غير أن ذلك نادر الاستخدام.

ثالثاً: التباين، الانحراف المعياري والعزوم :

1- التباين : لتفادي عيب الانحراف المتوسط وهو عدم الخضوع للعمليات الجبرية، فإنه يتم إيجاد متوسط مربعات انحرافات القيم عن وسطها الحسابي، ويسمى هذا المتوسط بالتباين.

أ- تباين البيانات غير المبوبة :

تعريف 5-4: إذا كانت لدينا البيانات: $x_1, x_2, x_3, \dots, x_n$ فإن تباينها يعرف كما يلي :

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N} \quad 5-5$$

مثال 5-5 : أوجد تباين بيانات المثال 5-3.

بتطبيق المعادلة 5-5 نجد :

$$\sigma^2 = \frac{\sum_{i=1}^{20} (x_i - \bar{x})^2}{N} = \frac{4080}{20} = 204$$

ب- تباين البيانات المبوبة :

تعريف 5-5: إذا كانت لدينا البيانات: $x_1, x_2, x_3, \dots, x_n$ تكراراتها على التوالي: $f_1, f_2, f_3, \dots, f_n$ فإن تباينها يعرف كما يلي :

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 f_i}{\sum_{i=1}^n f_i} \quad 6-5$$

i	x_i	f_i	$x_i f_i$	$(x_i - \bar{x})^2 f_i$
1	50	6	300	1536
2	60	4	240	144
3	70	5	350	80
4	80	3	240	588
5	90	1	90	576
6	100	1	100	1156
مج		20	1320	4080

جدول 3-5

$$\sigma^2 = \frac{4080}{20} = 204$$

نجد :

في حالة البيانات التي طول فئاتها أكبر من الصفر يتم استبدال x_i بمراكز الفئات وذلك في المعادلة 5-6.

إن التباين يأخذ في الحسبان جميع القيم وهو يخضع للعمليات الجبرية، إذ يمكن إيجاد قيمته أيضا عن طريق المعادلتين التاليتين :

$$\sigma^2 = \frac{\sum_{i=1}^n x_i^2}{N} - \bar{x}^2 \quad \text{في حالة البيانات غير}$$

المحبوبة.

$$\sigma^2 = \frac{\sum_{i=1}^n x_i^2 f_i}{\sum_{i=1}^n f_i} - \bar{x}^2 \quad \text{في حالة البيانات المبوبة.} \quad 8-5$$

أي أنه عبارة عن مربع الوسط التربيعي منقوصاً منه مربع الوسط الحسابي:

$$\sigma^2 = \mathbf{O}^2 - \mathbf{x}^2$$

و يمكن اثبات ذلك رياضيا بفك المعادلتين 5-5 و 6-5 على التوالي و إجراء بعض التعويضات.

عيب التباين الوحيد هو أن قيمته تكون كبيرة و وحدات قياسه تكون مربعة، لأنه يأخذ مربعات القيم في الحساب، الشيء الذي لا يجعله يعطي نظرة في تمام الوضوح حول مدى تشتت القيم، لذلك يتم في غالب الأحيان إستبداله بالانحراف المعياري كما هو معرف أدناه.

2- الانحراف المعياري : الانحراف المعياري هو الجذر التربيعي الموجب للوسط الحسابي لمربعات انحرافات القيم عن وسطها الحسابي، أي هو الجذر التربيعي الموجب للتباين، ويعرف رياضيا حسب طبيعة البيانات كما يلي :

1- الانحراف المعياري للبيانات غير المبوبة :

تعريف 5-6 : إذا كانت لدينا البيانات: $x_1, x_2, x_3, \dots, x_n$ فإن انحرافها المعياري يعطى كما يلي :

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N}} \quad 9-5$$

وعلى هذا يكون الانحراف المعياري لبيانات المثال 3-5 هو :

$$\sigma = \sqrt{204} = 14.28.10^3 \quad \text{دج}$$

ب- الانحراف المعياري للبيانات المبوبة :

تعريف 5-7: إذا كانت لدينا البيانات : $x_1, x_2, x_3, \dots, x_n$ ، تكراراتها على التوالي : $f_1, f_2, f_3, \dots, f_n$ فإن انحرافها المعياري يعطى بالمعادلة التالية:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 f_i}{\sum_{i=1}^n f_i}} \quad 10-5$$

مثال 5-7: أوجد الانحراف المعياري لبيانات الجدول 5-1.

بما أن الانحراف المعياري هو الجذر التربيعي الموجب للتباين، لذلك يكون:

$$\sigma = \sqrt{204} = 14.28.10^3 \text{ دج}$$

ج- خصائص الانحراف المعياري : من أهم خصائص الانحراف المعياري ما يلي:

* من أهم مميزاته، أنه يأخذ في الحسبان جميع القيم، كما أن قيمته صغيرة وبالتالي يمكن أن تعطي خلاصة واضحة عن مدى تباعد القيم، اذ كلما كانت هذه القيمة صغيرة دل ذلك على أن القيم ليست متباعدة عن الوسط الحسابي وبالتالي فهي أقل تشتتاً ووسطها الحسابي يمثلها تمثيلاً جيداً. و عموماً تعتبر القيم غير مشتتة إذا كان الانحراف المعياري يمثل أقل من 20 % من وسطها الحسابي.

* يمكن حساب الانحراف المعياري بالاعتماد على الوسط التربيعي كما هو معرف في مقاييس الترة المركزية باستخدام المعادلتين:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n x_i^2}{N} - \bar{x}^2} \quad 11-5 \quad \text{في حالة البيانات غير المبوبة.}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n x_i^2 f_i}{\sum_{i=1}^n f_i} - \bar{x}^2} \quad 12-5 \quad \text{في حالة البيانات المبوبة.}$$

و يمكن إثبات ذلك انطلاقاً من المعادلتين الأساسيتين 5-9 و 5-10.

* جذر متوسط مربعات انحرافات القيم عن وسطها الحسابي (الانحراف المعياري)، هو أصغر من جذر متوسط

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018
مربعات انحرافات تلك القيم عن أية قيمة أخرى مهما كانت
تختلف عن الوسط الحسابي لها أي:

$$\sqrt{\frac{(x_i - \bar{x})^2}{N}} < \sqrt{\frac{(x_i - Z)^2}{N}} \quad \forall \bar{x} \neq Z \quad 13-5$$

في حالة البيانات غير المبوبة. N.

$$\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 f_i}{\sum_{i=1}^n f_i}} < \sqrt{\frac{\sum_{i=1}^n (x_i - Z)^2 f_i}{\sum_{i=1}^n f_i}} \quad \forall \bar{x} \neq Z \quad 14-5$$

في حالة البيانات المبوبة.

وهذا مهما كان الوسط الحسابي للبيانات يختلف عن القيمة Z
و يتم إثبات ذلك بصفة مشاهة لإثبات الخاصية الرابعة من
خواص الوسط الحسابي كما هي واردة في الفصل الثاني.
* يمكن حساب الانحراف المعياري باستعمال وسط فرضي
Z عن طريق المعادلتين التاليتين :

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - Z)^2}{N} - (\bar{x} - Z)^2} \quad 15-5$$

في حالة البيانات غير المبوبة.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - Z)^2 f_i}{\sum_{i=1}^n f_i} - (\bar{x} - Z)^2} \quad 16-5$$

في حالة البيانات المبوبة.

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018
ويمكن إثبات ذلك بفك المعادلتين، إذ نجد هـا مساويتين
للمعادلتين 9-5 و 10-5 على التوالي.

* إذا كانت لدينا عيتين الأولى حجمها : N_1 وإنحرافها المعياري σ_1 ،
والثانية حجمها : N_2 وإنحرافها المعياري σ_2 ، ولهما نفس الوسط الحسابي،
فإن انحرافهما المعياري عند دمجهما يعطى على النحو التالي :

$$\sigma = \sqrt{\frac{N_1\sigma_1^2 + N_2\sigma_2^2}{N_1 + N_2}} \quad 17-5$$

ويمكن إثبات ذلك رياضيا أيضا.
ويظهر إذن أن الانحراف المعياري يوفر جميع شروط يول و هو
يعتبر بذلك أفضل مقياس من مقاييس التشتت، بحيث تعتمد
عليه معظم الدراسات الإحصائية.

3- العزم:

تعريف 4-8: إذا كانت لدينا القيم: $x_1, x_2, x_3, \dots, x_n$ حيث N عدد
القيم، فإن العبارة:

$$\bar{x}_q = \frac{x_1^q + x_2^q + x_3^q + \dots + x_n^q}{N} \quad 18-4$$

تسمى بالعزم من الدرجة q . (حيث q عدد طبيعي).
إذا كانت $q=1$ يكون العزم مساويا للوسط الحسابي.

تعريف 4-9: يسمى العزم من الدرجة q بالنسبة للوسط
الحسابي، بالعزم المركزي، ويعرف كما يلي :

$$m_q = \frac{\sum_{i=1}^n (x_i - \bar{x})^q}{N} \quad 19-4$$

في حالة البيانات غير المبوبة.

$$m_q = \frac{\sum_{i=1}^n (x_i - \bar{x})^q f_i}{\sum_{i=1}^n f_i} \quad 20-4$$

في حالة البيانات المبوبة.

إذا كان : $q=1$ يكون : $m_q=0$

إذا كان : $q=2$ يكون : $m_q=\sigma^2$ ، أي التباين.

ملاحظة: للتطبيق أنظر مثال 3-5 من الفصل الموالي.

رابعاً: الانحراف الربيعي:

تعريف 4-11: إذا كانت لدينا بيانات إحصائية رباعية ريعها الأول هو Q_1 وربيعها الثالث هو Q_3 ، فإن إنحرافها الربيعي يعطى كما يلي:

$$VQ = \frac{Q_3 - Q_1}{2} \quad 21-4$$

خامساً: معاملات التشتت النسبية: من عيوب مقاييس التشتت أن وحدات قياسها تأخذ نفس وحدات قياس المعلومات الأولية، لذلك لا يمكن إستخدامها في مقارنة مدى تشتت ظاهرتين مختلفتين أو أكثر، فعلى سبيل المثال، التشتت في ظاهرة النفقات في الجزائر يقاس بالدينار الجزائري بينما في دولة أخرى يقاس بعملية تلك الدولة ، لذلك لا يمكن المقارنة بين تشتت الظاهرتين، الشيء الذي تطلب ضرورة إيجاد مقاييس نسبية تسمح بالمقارنة ومنها معامل الاختلاف أو معامل التغير والانحراف الربيعي النسبي.

1- معامل الاختلاف:

تعريف 4-10: يعرف معامل الاختلاف بأنه النسبة المئوية للانحراف المعياري على الوسط الحسابي، أي :

$$CV = \frac{\sigma}{x} \times 100 \quad 22-4$$

بما أن وحدات قياس كل من الانحراف المعياري و الوسط الحسابي واحدة، لذلك تختزل من بسط ومقام المعادلة 22-4، ويكون بذلك CV قيمة نسبية.

2- الانحراف الربيعي النسبي : وهو يستخدم في تحديد مدى تماثل التوزيعات، ويعطى كما يلي :

$$VQP = \frac{Q_3 - Q_1}{Q_2} \quad 23-4$$

ملاحظة: العلاقة بين بعض مقاييس التشتت: إذا كانت لدينا بيانات تكرارية ذات إتواء ضعيف (انظر الفصل الموالي)، فإنه ثبت تجريبيا صحة العلاقات التالية:

الانحراف المتوسط = أربعة أخماس الانحراف المعياري،

أي:

$$e = \frac{4}{5} \sigma \quad 24-4$$

الانحراف الربيعي = ثلثي الانحراف المعياري، أي :

$$VQ = \frac{2}{3} \sigma \quad 25-4$$

تمارين.

تمرين 1: عرف التشتت وحدد مقاييسه . ماهو المقياس الأكثر إستخداما؟ وضح لماذا؟

تمرين 2: البيانات التالية تظهر نتائج أحد الأقسام في مقياس الإحصاء:

6	2	10	12	4	6	10	4	8	2
2	8	12	4	12	2	4	6	10	4

المطلوب: أوجد المقاييس التالية : - الانحراف المتوسط . - الانحراف المعياري . - معامل الاختلاف - الانحراف الربيعي - الانحراف الربيعي النسبي - العزم المركزي من الدرجة 2 و 3.

تمرين 3: بوب بيانات التمرين 2 ، وأجب على كل مطالبيه.

تمرين 4: مصنع ينتج نوعين من العجلات ، النوع الأول متوسط المسافة التي يهتلك فيها هي : 10000 كلم بإنحراف معياري يبلغ 2000 كلم، أما النوع الثاني فمتوسط المسافة التي يهتلك فيها هي : 11000 كلم بإنحراف معياري يقدر بـ: 1000 كلم ، هل يمكن القول أن النوع الثاني أفضل من النوع الأول.

تمرين 5: الجدول التالي يبين توزيع عينة من مساكن أحد الأحياء حسب عدد الغرف.

7	6	5	4	3	2	1	عدد الغرف x_i
3	5	12	30	25	20	5	عدد المساكن f_i

المطلوب :

- 1- أوجد كل مقاييس التشتت.
- 2- أوجد معامل الاختلاف.
- 3- أوجد العزم المركزي من الدرجة 2 و 4

تمرين 6: البيانات التالية تظهر توزيع عمال مصنعين متشابهين الأول في الجزائر و الثاني في واشنطن حسب الأجور الشهرية التي يتقاضونها.

توزيع الأجور الأسبوعية للمصنع 1 بآلاف الدينارات

الأجور	4-2	6-4	8-6	10-8	12-10	14-12	16-14	18-16
عدد العمال	10	35	40	64	35	10	5	1

توزيع الأجور الأسبوعية للمصنع 2 بآلاف الدولارات

الأجور	4-2	6-4	8-6	10-8	12-10	14-12	16-14	8-16
عدد العمال	20	35	50	39	30	20	4	2

المطلوب: 1- أوجد كل مقاييس التشتت للمصنعين. 2- أي المصنعين أفضل توزيعا. إشرح.

تمرين 7: أثبت أنه يمكن كتابة:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n x_i^2}{N} - \bar{x}^2} \quad \text{أ- الانحراف المعياري للبيانات غير المبوبة على النحو:}$$

$$\sigma^2 = \frac{\sum_{i=1}^n x_i^2 f_i}{\sum_{i=1}^n f_i} - \bar{x}^2 \quad \text{ب- تباين البيانات المبوبة على النحو:}$$

ج- الانحراف المعياري للبيانات غير المبوبة على النحو:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - Z)^2}{N} - (\bar{x} - Z)^2}$$

حيث Z قيمة فرضية.

د- الإنحراف المعياري للبيانات المبوبة على النحو:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - Z)^2 f_i}{\sum_{i=1}^n f_i} - (\bar{x} - Z)^2}$$

حيث Z قيمة فرضية.

تمرين 8: إذا كانت لدينا بيانات ذات إتواء ضعيف ، إنحرافها المتوسط هو 8 ، أوجد تباينها.

تمرين 9: إذا كانت لدينا عيتين الأولى حجمها N_1 وإنحرافها المعياري σ_1 ، والثانية حجمها N_2 وإنحرافها المعياري σ_2 ، ولهما نفس الوسط الحسابي ، أثبت أن تباينهما عند دمجها يعطى على النحو التالي:

$$\sigma^2 = \frac{N_1 \sigma_1^2 + N_2 \sigma_2^2}{N_1 + N_2}$$

تمرين 10: أثبت أن التباين عبارة عن مربع الوسط التربيعي منقوصاً منه مربع

$$\sigma^2 = Q^2 - \bar{x}^2$$

الوسط الحسابي أي :

الفصل السادس

أشكال التوزيعات التكرارية.

عند رسم المنحني التكراري للبيانات نصادف عدة أنواع من الأشكال، كل شكل يوحي بطبيعة معينة لتوزيع تلك البيانات، الشيء الذي يجعل مقاييس الترة المركزية ومقاييس التشتت وحدها لا تكفي لتحليلها، ومن الأشكال التي يمكن مصادفتها ما يصطلح عليه بما يلي:

* التماثل التام. * الإلتواء. * التفرطح (الإنسباط).

أولاً: التماثل التام: قد يكون المنحني التكراري متماثلاً، بحيث يكون الشق الأيسر للشاقول الواصل بين قمة التوزيع والمحور الأفقي للمعلم، مماثلاً تماماً للشق الأيمن له، وفي هذه الحالة تكون نقطة تقاطع الشاقول مع المحور الأفقي للمعلم هي النقطة المركزية التامة للتوزيع، وتسمى بنقطة التماثل، وتكون مساوية للوسط الحسابي وللوسط والمنوال، وتكون بالتالي ممثلة تمثيلاً جيداً للتوزيع.

تعريف 6-1: إذا كانت لدينا البيانات: $x_1, x_2, x_3, \dots, x_n$ تكراراتها على التوالي: $f_1, f_2, f_3, \dots, f_n$ ، نقول أن التوزيع التكراري لهذه البيانات تام التماثل إذا توافرت الشروط التالية:

$$\bar{x} = M_e = M_o \quad * \quad 6-1$$

* تكرارات الفئات التي تقل عن القيمة المركزية مساوية للتكرارات التي تزيد عنها.

* الفرق بين كل فئة و الفئة التي تليها متساو، إذا

كان $L=0$ ، وأطوال الفئات متساو إذا كان $L > 0$

مثال 6-1: أثبت تماثل البيانات التالية :

بواسطة

153 باحث الاقتصادي

Economic Research Gate

i	x_i	f_i
1	10	2
2	12	3
3	14	5
4	16	3
5	18	2
مج		15

جدول 6-1

الإجابة: لإثبات ذلك لابد أن نتأكد من شروط التماثل التام :
الشرط الأول:
 * الوسط الحسابي :

$$\bar{x} = \frac{\sum_{i=1}^5 x_i f_i}{\sum_{i=1}^5 f_i} = \frac{210}{15} = 14$$

* الوسيط :

$$c = \frac{\sum_{i=1}^5 f_i + 1}{2} = \frac{16}{2} = 8$$

ترتيب الوسيط هو :

أي أن قيمة الوسيط هي التي ترتيبها 8 ومنه فإن :

$$M_e = 14$$

* المنوال : وهو الفئة الأكثر تكرارا، ومنه يكون :

$$M_o = 14$$

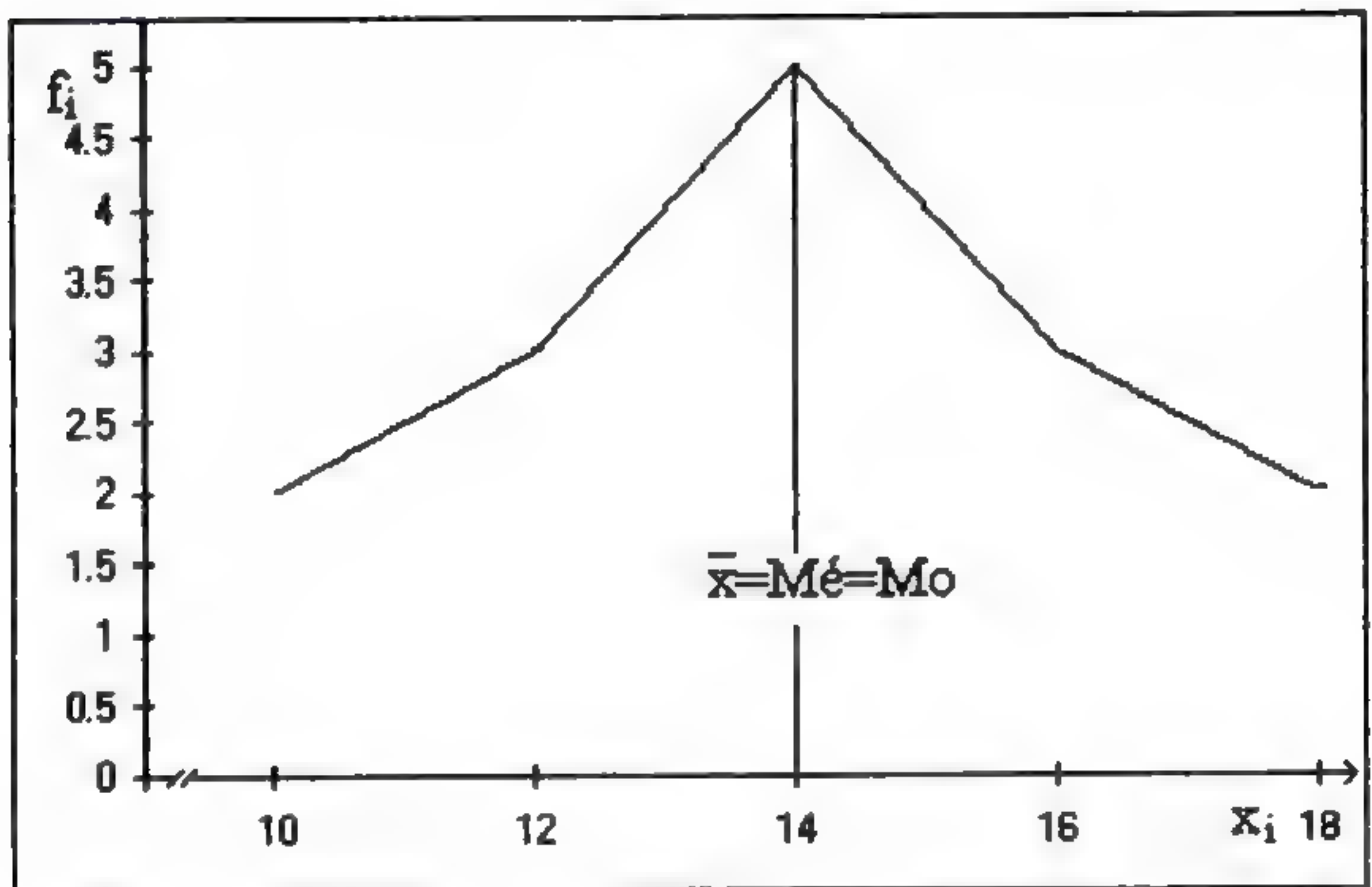
وعليه فإن الشرط الأول محقق، أي :

$$\bar{x} = M_e = M_o = 14$$

الشرط الثاني: التكرارات التي تقل عن الوسط الحسابي،
 تساوي تلك التي تزيد عنه، وهذا الشرط محقق أيضا.

الشرط الثالث : الفرق بين كل فئة والفئة التي تليها متساو
 ويساوي 2.

وبالتالي فإن التوزيع المشار اليه تام التماثل. وذلك ما يوضحه الشكل الموالي :



شكل 6-1

إن التوزيعات التامة التماثل نادرة المصادفة في الحياة العملية، بينما التوزيعات غير المتماثلة فهي كثيرة المصادفة، وتعرف بالتوزيعات الملتوية.

ثانياً: الالتواء: الحالة الثانية التي يمكن مصادفتها عند رسم المنحنى التكراري، هي عدم التماثل، وهي الحالة التي لا تتوافر فيها شروط التماثل كما هي موضحة أعلاه، و عنها نقول أن التوزيع ملتو إما الى اليمين أو الى اليسار.

كما سبقت الإشارة فانه عند دراسة أية ظاهرة من الظواهر، لا يكفي تحليل ميل عناصر الظاهرة نحو القيمة المركزية، ولا

تباعده تلك القيم عنها، اذ لابد من دراسة درجة ميل قيم الظاهرة نحو قيمة معينة تزيد أو تقل عن قيمتها المركزية، أي لابد من دراسة إلتواءها.

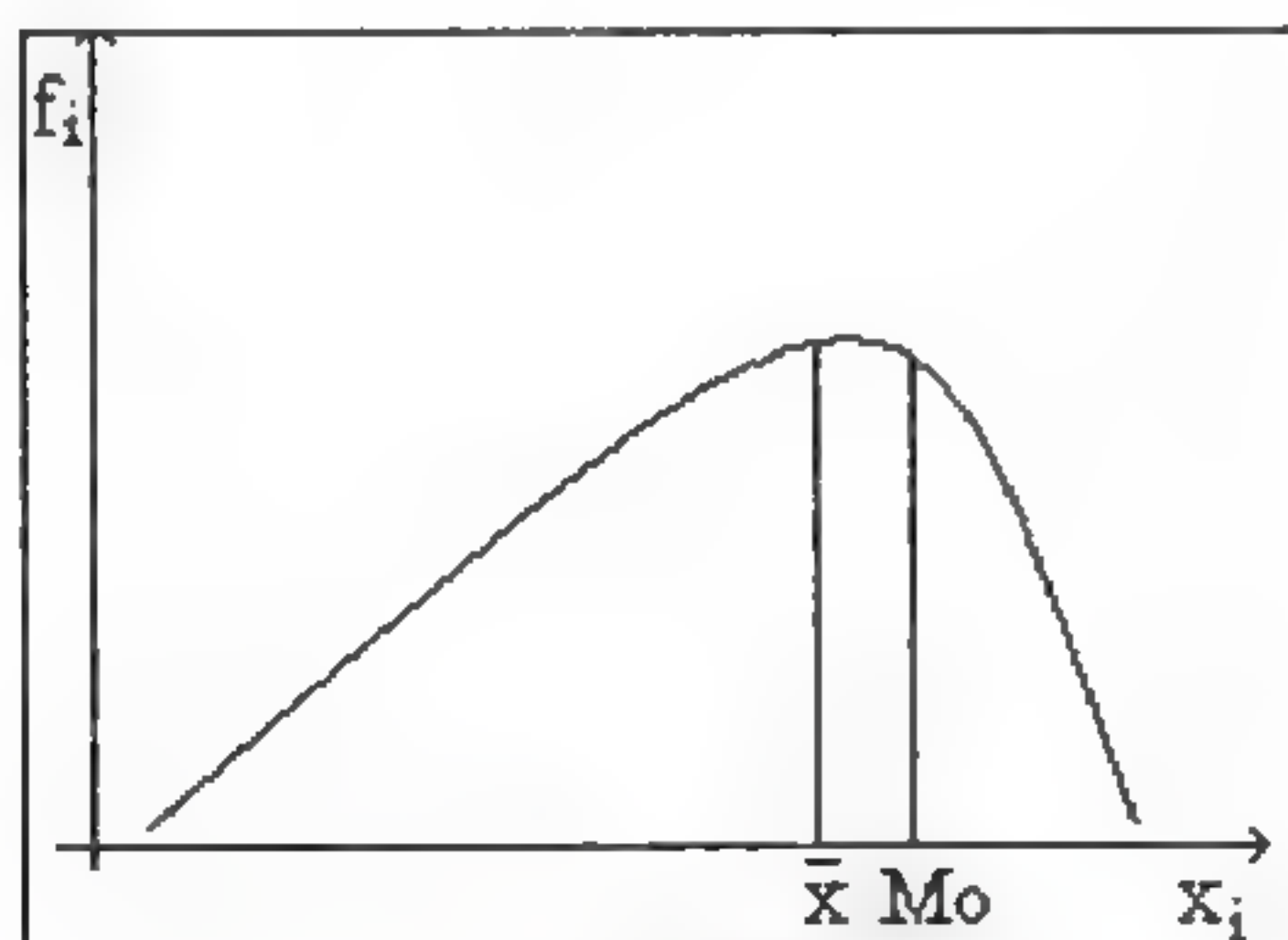
تعريف 2-6: يقصد بالإلتواء إنعدام التماثل في توزيع قيم الظاهرة حول قيمتها المركزية (الوسط الحسابي)، وفيه تنتفي شروط التماثل التام، ويقاس الإلتواء بعدة مقاييس منها ما يلي :

1- قيمة الإلتواء: لمعرفة قيمة إلتواء ظاهرة ما يتم استخدام المعادلة التالية:

$$VA = \bar{x} - Mo \quad 2-6$$

أي أن قيمة الإلتواء عبارة عن الوسط الحسابي منقوصاً منه المنوال.

يكون التوزيع سالب الإلتواء إذا كان المنحنى التكراري ممتداً الى اليسار، كما في الشكل 2-6.

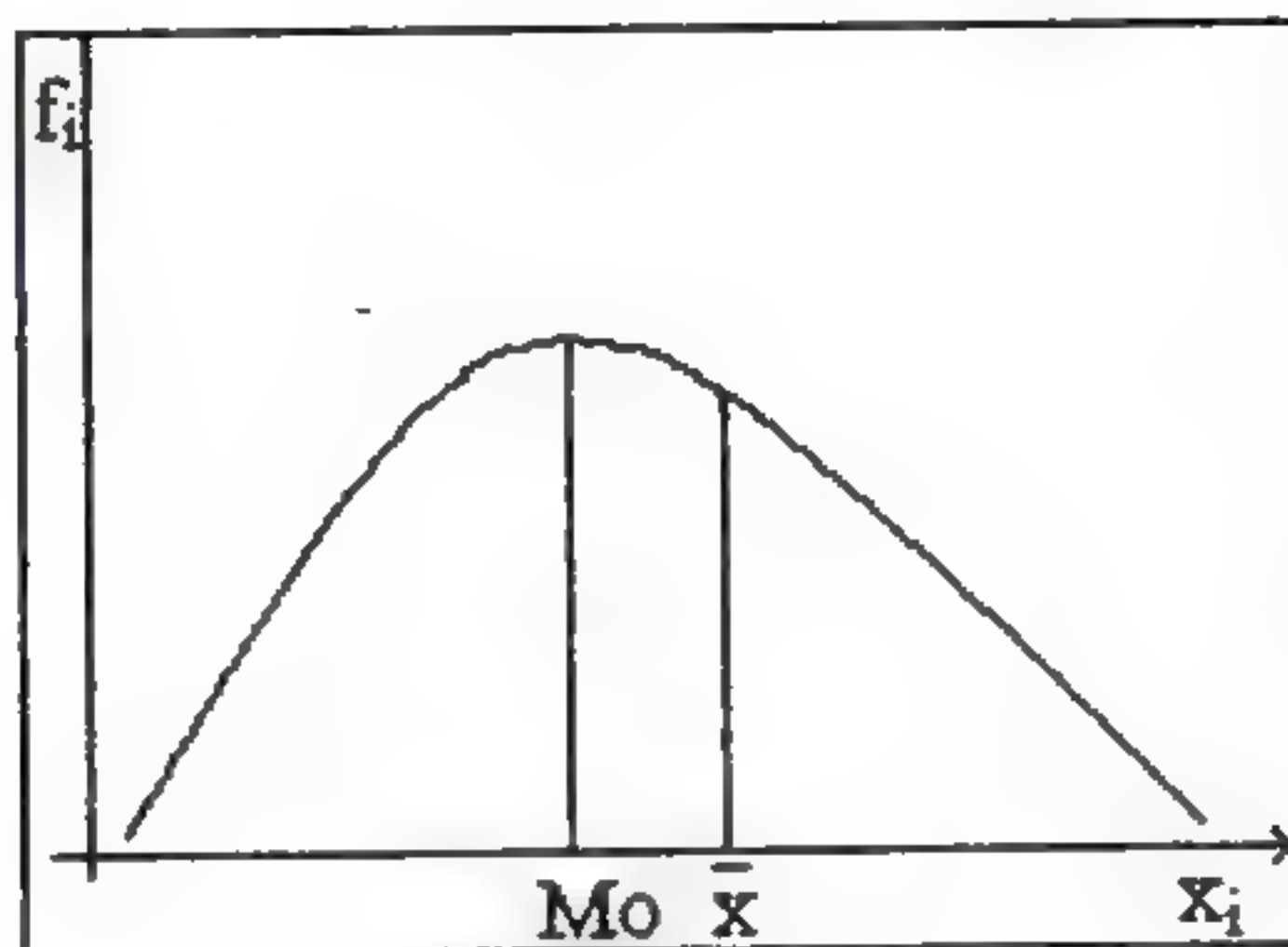


شكل 2-6

$$\bar{x} < Mo$$

3-6

ويكون التوزيع موجب الالتواء إذا كان المنحنى التكراري ممتدا الى اليمين، كما في الشكل 3-6.



شكل 3-6

$$\bar{x} > Mo$$

4-6

2- معامل بيرسون للالتواء: إن قيمة الالتواء كما وردت في المعادلة 2-6، لا تسمح بمقارنة التواء ظاهرتين مختلفتين أو أكثر نتيجة لاختلاف وحدات القياس، لذلك يتم استخدام ما يسمى بمعامل بيرسون للالتواء الذي يعطى بالمعادلة التالية:

$$CA = \frac{\bar{x} - Mo}{\sigma} \quad 5-6$$

أي قيمة الالتواء مقسومة على الانحراف المعياري.
من المعادلة 4-4 نستنتج :

$$\bar{x} - Mo = 3(\bar{x} - Me) \quad 6-6$$

وذلك في حالة الالتواء، لذلك فإن معامل بيرسون للالتواء يكتب أيضا بالصيغة :

$$CA = \frac{3(\bar{x} - Mé)}{\sigma}$$

7-6

ويكون معامل الالتواء محصورا بين $1+$ و $1-$.
إذا كان موجبا دل على أن التوزيع يمتد الى اليمين، أما إذا كان سالبا دل على أن التوزيع يمتد الى اليسار.
وتم تقسيم قيمة الالتواء على الانحراف المعياري كما سبقت الإشارة لإختزال وحدات القياس، والتمكن بالتالي من مقارنة إلتواء ظاهرتين مختلفتين.

مثال 6-2: الجدولين التاليين يظهران توزيع مجموعة من العمال حسب الأجور الشهرية التي يتقاضونها، بآلاف الدينارات في المؤسستين أ و ب.

توزيع أجور عمال
المؤسسة ب

عدد العمال	الأجر	i
2	160	1
4	200	2
5	250	3
8	270	4
1	300	5
20	مج	

جدول 6-3

توزيع أجور عمال
المؤسسة أ

عدد العمال	الأجر	i
1	100	1
8	110	2
5	130	3
4	170	4
2	190	5
20	مج	

جدول 6-2

المطلوب:

- 1- أوجد قيمة الالتواء لكل توزيع.
- 2- أوجد معامل بيرسون للالتواء.
- 3- قارن بين إلتواء التوزيعين، أثبت النتيجة أيضا بالرسم.

الجواب:

1- لإيجاد قيمة الإلتواء، نوجد الوسط الحسابي و المنوال لكلا التوزيعين.

الوسط الحسابي للتوزيع الأول: (نجده باستخدام المعادلة رقم (4-3)

$$\bar{x}_1 = \frac{2690}{20} = 134.5$$

الوسط الحسابي للأجور في التوزيع الأول هو : 134.5 ألف دينار.

الوسط الحسابي للتوزيع الثاني: (نجده باستخدام المعادلة رقم (4-4

$$\bar{x}_2 = \frac{4830}{20} = 241.5$$

الوسط الحسابي للأجور في التوزيع الثاني هو : 241.5 ألف دينار.

منوال التوزيع الأول: $Mo_1 = 110$ ألف دينار.

منوال التوزيع الثاني: $Mo_2 = 270$ ألف دينار.

ومنه فإن قيمة إلتواء التوزيع الأول هي :

$$VA_1 = \bar{x}_1 - Mo_1 = 134.5 - 110 = 24.5$$

أي أن قيمة إلتواء التوزيع الأول هي: 24.5 ألف دينار، ويدل

ذلك على أن بيانات المؤسسة الأولى ذات إلتواء شديد الى

اليمين.

أما قيمة إلتواء التوزيع الثاني :

$$VA_2 = \bar{x}_2 - Mo_2 = 241.5 - 270 = -28.5$$

قيمة إلتواء التوزيع الثاني هي : -28.5 ألف دينار، ويدل ذلك على أن بيانات المؤسسة الثانية ذات إلتواء شديد الى اليسار.

واضح أن قيمة الإلتواء أخذت نفس وحدة قياس المعلومات الأولية، لكون كل من الوسط الحسابي والمنوال يأخذ نفس وحدة قياس المعلومات الأولية.

2- لإيجاد معامل بيرسون للإلتواء يتم إيجاد الانحراف المعياري للتوزيعين:

الانحراف المعياري للتوزيع الأول: (نجده باستخدام المعادلة رقم : 5-10)

$$\sigma_1 = \sqrt{\frac{17295}{20}} = 29.41$$

الانحراف المعياري للتوزيع الثاني: (نجده باستخدام المعادلة رقم : 5-10)

$$\sigma_2 = \sqrt{\frac{30455}{20}} = 39.02$$

و منه فان معامل بيرسون لإلتواء التوزيع الأول هو:

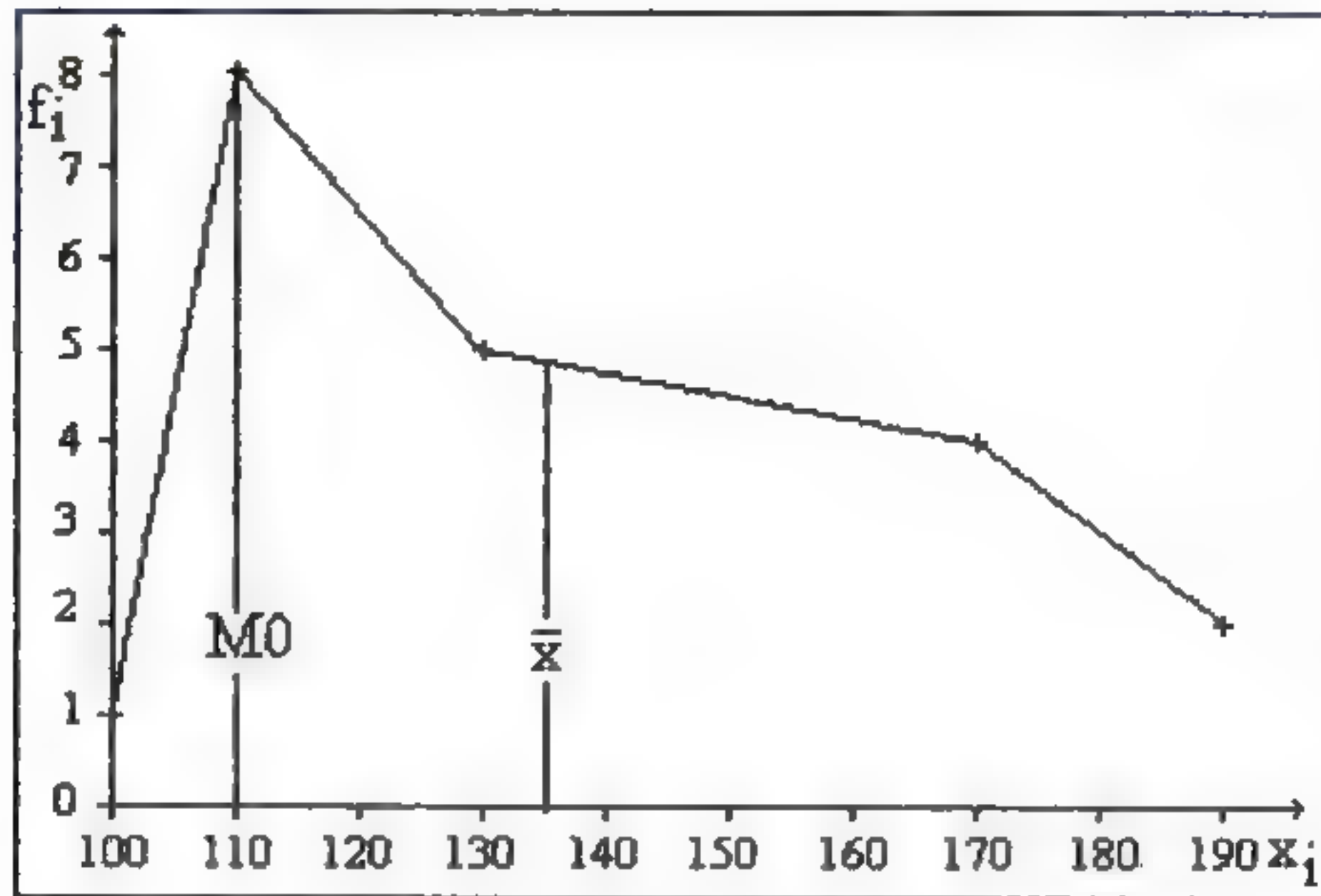
$$CA_1 = \frac{134.5 - 110}{29.41} = 0.83$$

أما معامل بيرسون لإلتواء التوزيع الثاني فهو :

$$CA_2 = \frac{241.5 - 270}{39.02} = -0.73$$

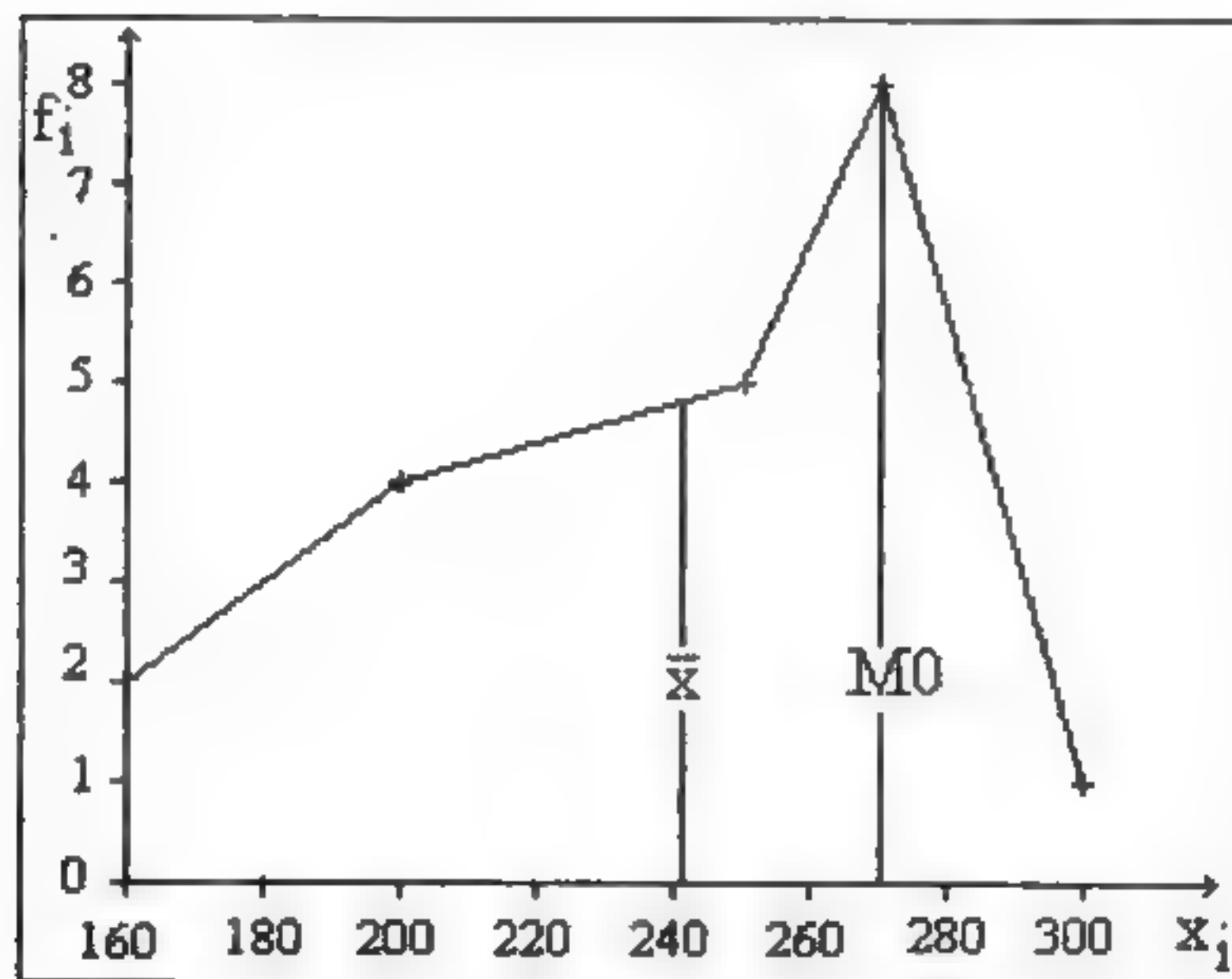
ومنه نلاحظ أن التوزيع الثاني ملتو الى اليسار، بينما التوزيع الأول ملتو الى اليمين، وهو أكثر إلتواء من التوزيع الثاني، وذلك ما يوضحه الشكلين التاليين:

المنحنى التكراري للتوزيع الأول



شكل 4-6

المنحنى التكراري للتوزيع الثاني.



شكل 5-6

وهناك مقاييس أخرى للإلتواء أقل إستخداما، منها معامل الإلتواء الربيعي (معامل يول) وهو يعتمد على الربيعيات، ومعامل الإلتواء العزمي (معامل بيرسون)، وهو يعتمد على العزوم، ويتم استخدامهما خاصة في حالة وجود أكثر من منوال واحد للبيانات.

3-معامل الإلتواء الربيعي (معامل يول) :

تعريفه 3-6: إذا كانت لدينا مجموعة من البيانات ربيعياتها الأول و الثاني (الوسيط) والثالث هي على التوالي: Q_1, Q_2, Q_3 ، فإن معامل الإلتواء الربيعي لها يعطى بالمعادلة التالية:

$$CAQ = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)} \quad 8-6$$

تكون قيمة هذا المعامل محصورة بين $1+$ و $1-$. إذا كانت قيمته قريبة من الصفر يعني ذلك أن التوزيع قريبا من التناظر، وتدل إشارته إلى إتجاه الإلتواء إلى اليمين أو إلى اليسار.

4-معامل الإلتواء العزمي (معامل بيرسون):

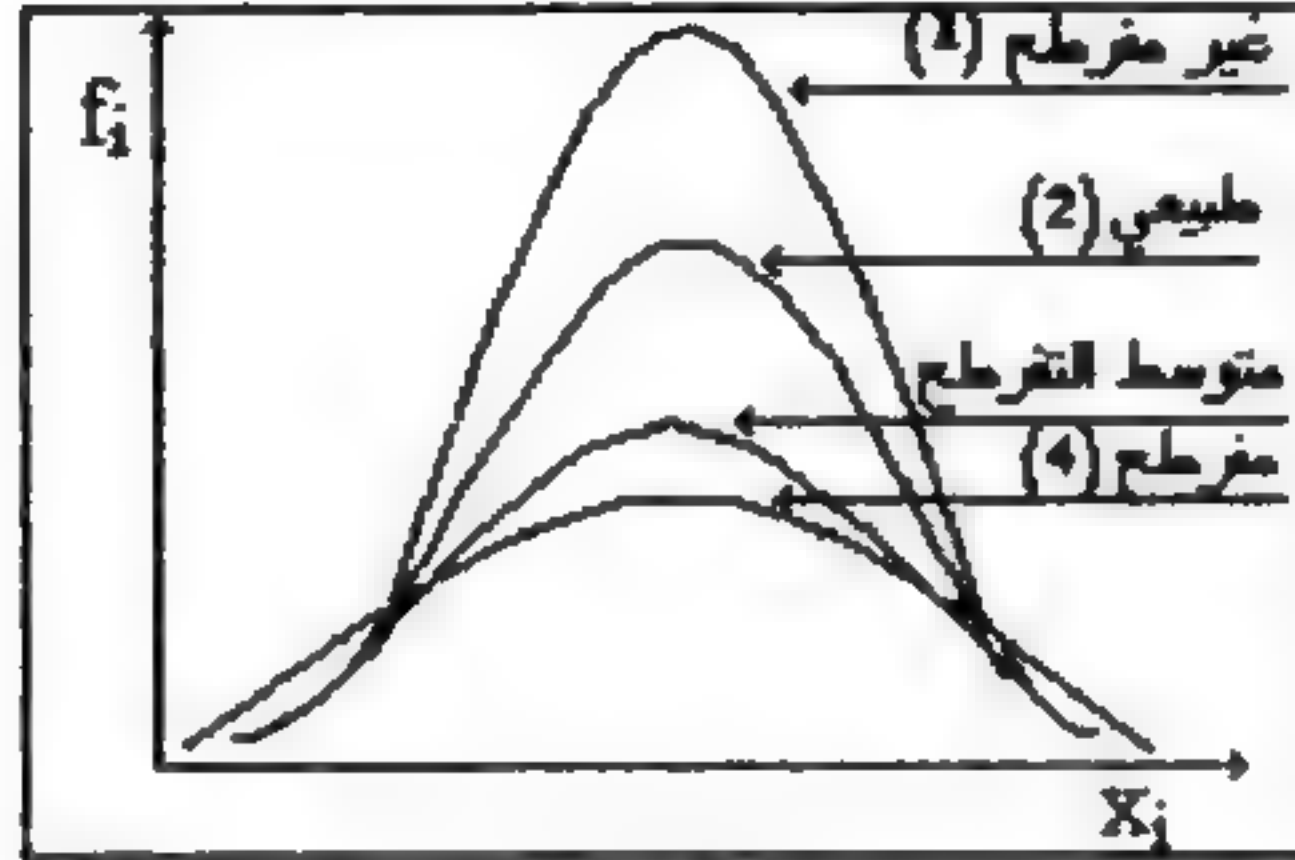
تعريفه 4-6: معامل الإلتواء العزمي هو النسبة بين العزم الثالث ومكعب الانحراف المعياري لتلك البيانات، أي :

$$CAm = \frac{m^3}{\sigma^3} \quad 9-6$$

يستخدم هذا المعامل إذا كان التوزيع وحيد المنوال، وتنحصر قيمته أيضا بين $1+$ و $1-$ ، وتدل الإشارة على اتجاه التوزيع. كما سبقت الإشارة فإن هذه المعاملات تستخدم لمعرفة أشكال التوزيعات، كما تستخدم أيضا لمقارنة أشكال التوزيعات لعدة ظواهر، غير أنه في هذه الحالة ينبغي استعمال نفس القانون.

مثال: التفرطح (الإنساط): كما سبق وأن رأينا، فقد يكون المنحنى التكراري متمائلا أو ملتويا إلى اليمين أو إلى اليسار،

وفي كلا الحالتين قد يكون المنحنى عالي القمة (غير مفطح) أو منبسط القمة (متوسط التفرطح أو مفطح)، وتقاس درجة علو المنحنى بما يسمى بمعامل التفرطح، والشكل التالي يظهر مختلف أنواع المنحنيات التكرارية الممكن مصادفتها:



شكل 6-6

المنحنى (1) (غير مفطح - عالي القمة - مدبب)، يدل المنحنى على شدة تركيز القيم حول مقاييسها المركزية، على عكس المنحنى الرابع، وهو منحنى مفطح، حيث تكون القيم متباعدة عن بعضها البعض، وغير متمركزة حول وسطها الحسابي ويكون المنحنى منبسط القمة، أما المنحنى الثاني فهو منحنى الحالة العادية (الطبيعية)، ويسمى بمنحنى التوزيع الطبيعي و يشبه الجرس، وفيه تتوزع تكراراته كما يلي:

68.27 % من التكرارات محصورة بين : $\bar{x} - \sigma$ و $\bar{x} + \sigma$

95.45 % من التكرارات محصورة بين : $\bar{x} - 2\sigma$ و $\bar{x} + 2\sigma$

99.73 % من التكرارات محصورة بين : $\bar{x} - 3\sigma$ و $\bar{x} + 3\sigma$

حيث σ : الانحراف المعياري.

و يقاس تفرطح المنحنيات التكرارية عن طريق عدة مقاييس منها :

1- معامل فيشر: الذي يعطى بالمعادلة التالية :

$$CF = \frac{m_4}{\sigma^4}$$

10-6

حيث: m_4 العزم من الدرجة الرابعة، σ^4 الانحراف المعياري مرفوع الى القوة 4. إذا كان معامل فيشر موجبا دل ذلك على أن التوزيع أقل تفرطحا من التوزيع الطبيعي، وإذا كان سالبا دل ذلك على أن التوزيع أكثر تفرطحا، أي مديبا أكثر من التوزيع الطبيعي.

2- معامل كيلي: يمكن حساب معامل التفرطح بطريقة كيلي كما يلي :

$$CK = \frac{Q_3 - Q_1}{D_9 - D_1} \quad 11-6$$

مثال 6-3: أوجد معامل الإلتواء العزمي ومعامل التفرطح لفisher، للبيانات التالية:

i	x_i	f_i
1	7	10
2	10	12
3	13	14
4	16	8
5	19	6
6	22	4

جدول 6-4

الاجابة : معامل الإلتواء العزمي يعطى بالمعادلة 6-8 بينما معامل التفرطح يعطى بالمعادلة 6-10، ويتم ايجادهما بمساعدة الجدول التالي :

i	x_i	f_i	$x_i f_i$	$(x_i - \bar{x})^2 f_i$	$(x_i - \bar{x})^3 f_i$	$(x_i - \bar{x})^4 f_i$
1	7	10	70	360	-2160	12960
2	10	12	120	108	-324	972
3	13	14	182	00	00	00
4	16	8	128	72	216	648
5	19	6	114	216	1296	7776
6	22	4	88	324	2916	26244
مج		54	702	1080	1944	48600

جدول 6-5

$$\bar{x} = \frac{702}{54} = 13$$

الوسط الحسابي :

$$\sigma = \sqrt{\frac{1080}{54}} = 4.47$$

الانحراف المعياري:

$$m_3 = \frac{1944}{54} = 36$$

العزم من الدرجة الثالثة:

$$m_4 = \frac{48600}{54} = 900$$

العزم من الدرجة الرابعة:

و منه نجد :

1- معامل الالتواء العزمي : (بتطبيق المعادلة رقم 6-9) :

$$CAm = \frac{36}{(4.47)^3} = 0.4$$

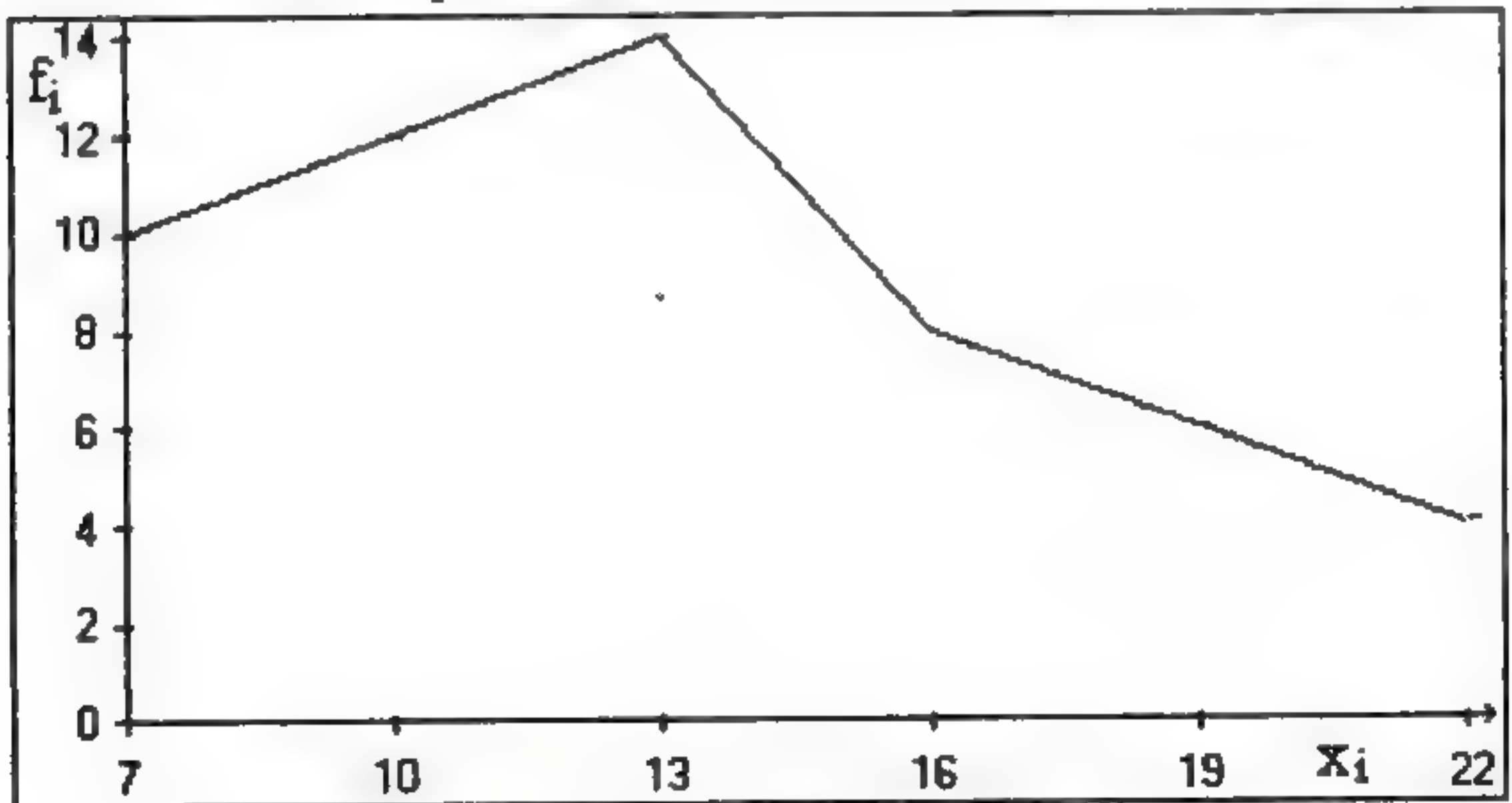
تدل النتيجة على أن التوزيع ملتو قليلا الى اليمين.

$$CF = \frac{900}{(4.47)^4} = 2.25$$

2- معامل التفرطح:

و بالتالي يكون التوزيع غير مفرطح.

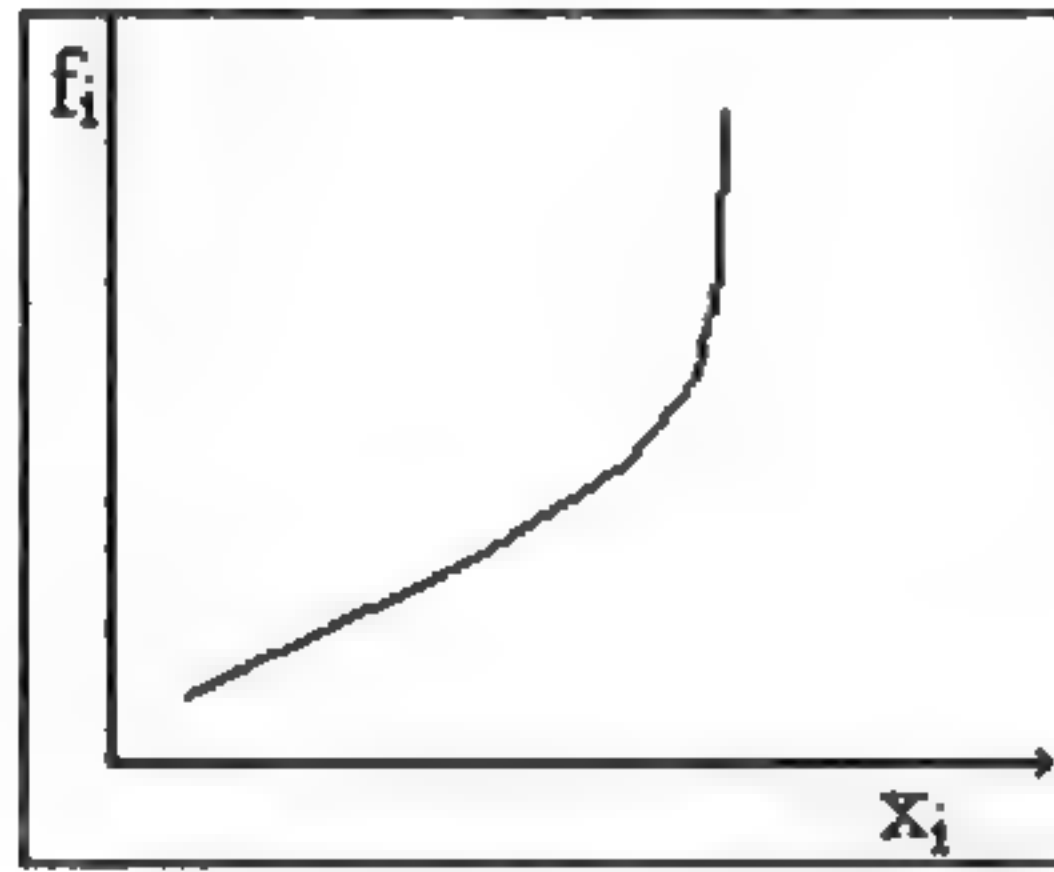
و يمكن ملاحظة ذلك من خلال الشكل الموالي.



شكل 6-6

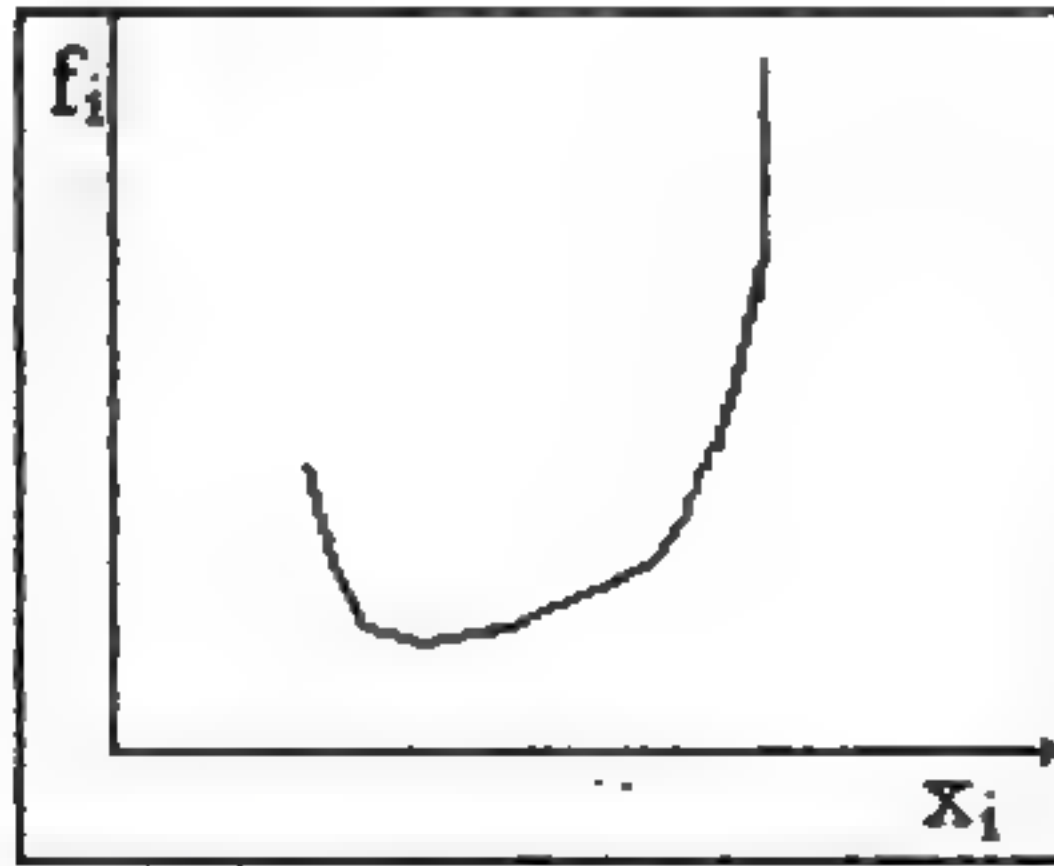
رابعاً : أشكال أخرى : إضافة الى أشكال المنحنيات التي تطرقنا لها لحد الآن هناك أشكال أخرى للتوزيعات التكرارية، غير أنها قليلة المصادفة في الحياة العملية، منها ما يلي :

1- المنحنى الرائي : يشبه شكله حرف ر، لذلك سمي بهذا الاسم، وهو كما في الشكل التالي :



شكل 6-7

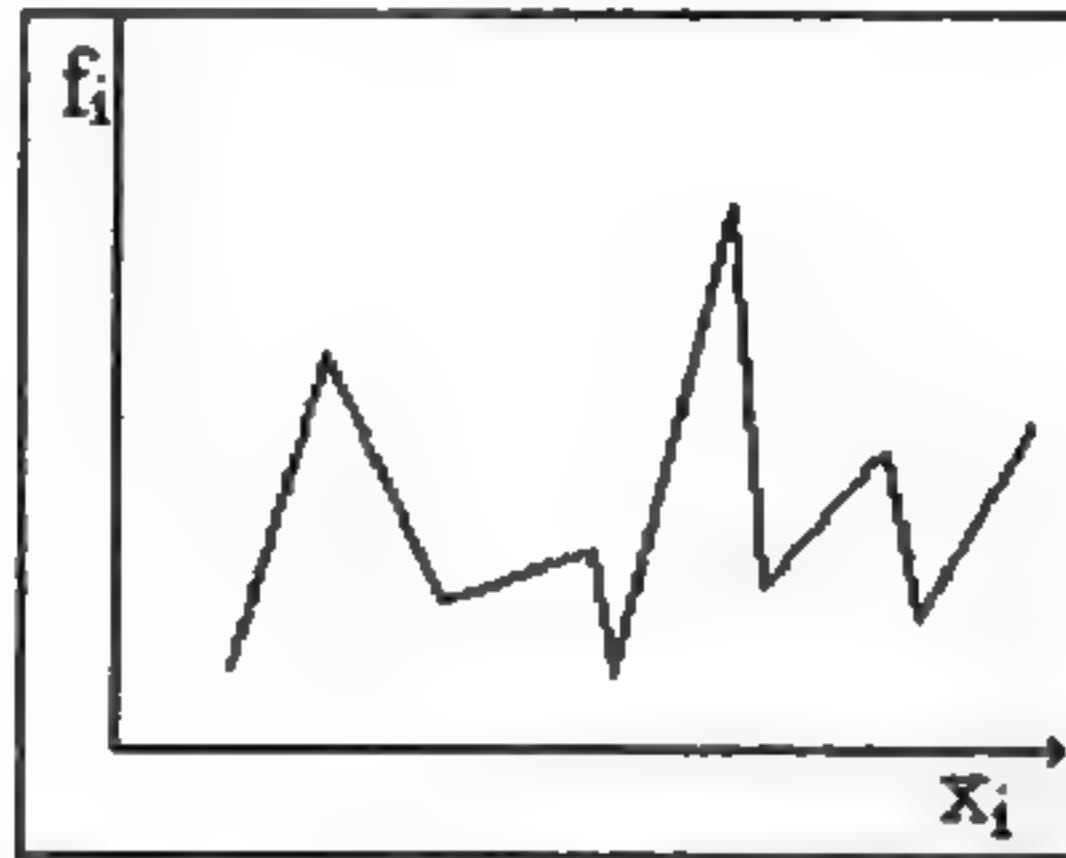
2- المنحنى اللامي : يشبه حرف اللام، و يوضحه الشكل التالي :



شكل 6-8

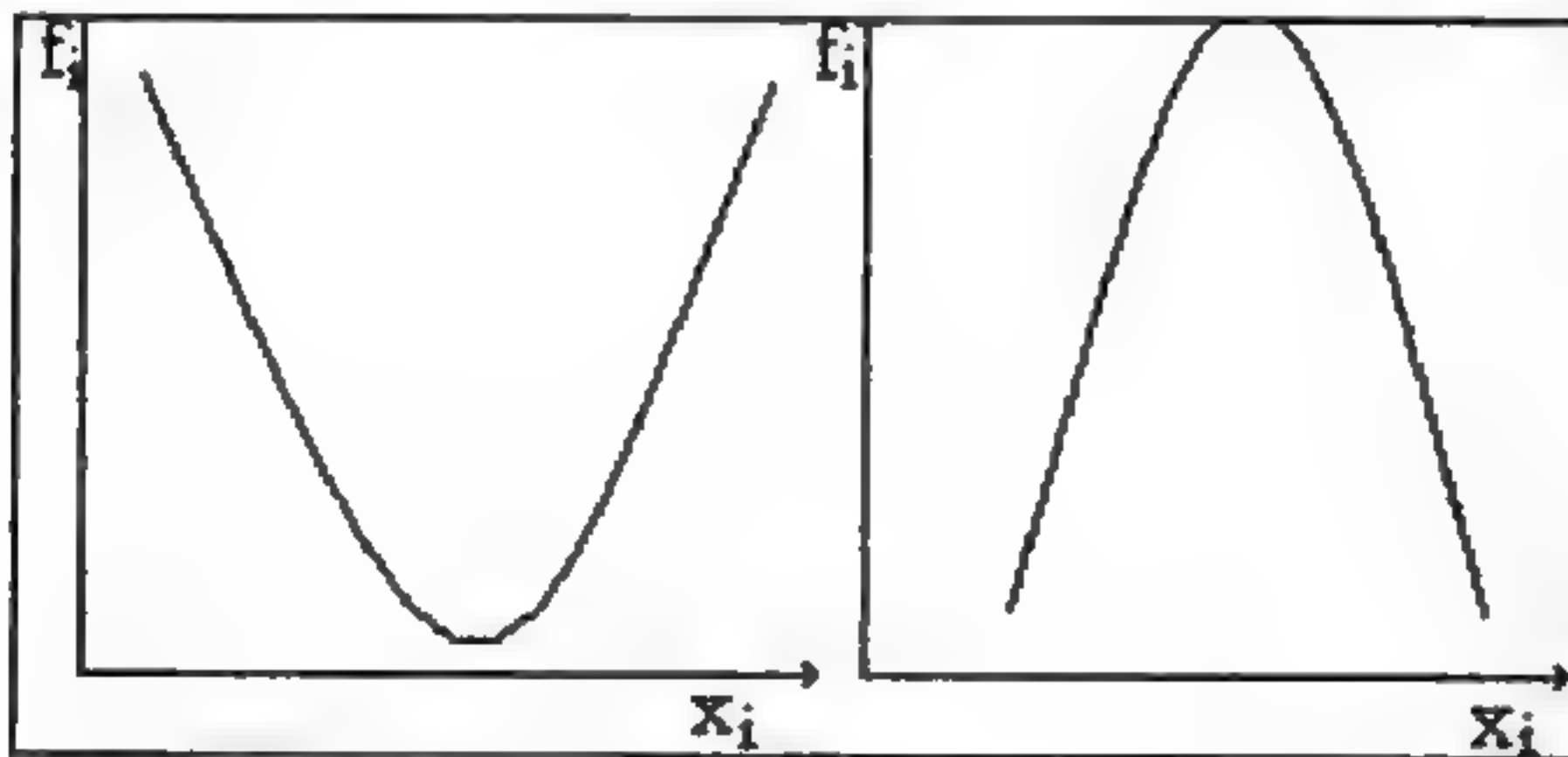
يمكن أن يكون المنحنى الرائي أو اللامي مقلوبين.

3- **منحنى متعدد القمم** : وهو المنحنى الذي تكون فيه بعض التكرارات بارزة، أو الذي يكون توزيعه متعدد المنوالات، ويكون شكله كما يلي :



شكل 6-9

4- **منحنيات القطع المكافئ** : وهي التي يأخذ منحناها أحد الشكلين التاليين :



شكل 6-11

شكل 6-10

يمكن مصادفة الأشكال 6-9، 6-10، 6-11 خاصة في دراسة السلاسل الزمنية.

أمثلة وتمارين

- تمرين 1:** حدد مفهومي التماثل التام والالتواء.
- تمرين 2:** ماهي الشروط التي يجب توافرها ليكون أي توزيع خاضع للتوزيع الطبيعي.
- تمرين 3:** أوجد معامل إلتواء بيرسون لبيانات التمرين 6 من سلسلة تمارين الفصل الخامس. هل هذا التوزيع ملتو؟ اشرح. قدم البيانات من خلال منحى تكراري. هل النتائج متطابقة؟
- أوجد معامل الإلتواء الربيعي ومعامل الإلتواء العزمي لنفس البيانات. ماذا تستنتج؟
- تمرين 4:** أوجد معامل التفرطح لبيانات التمرين 6 من سلسلة تمارين الفصل الخامس. ماذا تستنتج؟
- تمرين 5:** من بيانات التمرين 6 من سلسلة تمارين الفصل الخامس أوجد معامل الإلتواء العزمي لكل توزيع. قارن بين التوزيعين من خلال النتائج. ارسم منحنيهما التكراريين. هل النتائج متطابقة.
- تمرين 6:** أجب على نفس أسئلة التمرين 5، للتوزيعات التالية.

x_j	f_j
5	5
10	6
15	10
20	20
25	10
30	6
35	5

x_j	f_j
5	5
10	6
15	10
20	15
25	6
30	2
35	2

x_j	f_j
5	1
10	5
15	9
20	15
25	14
30	12
35	8

الفصل السابع

الإنحدار والارتباط

خلال الفصول السابقة، تعرضنا للظواهر ذات المتغير الواحد وأدوات تحليلها الأساسية، غير أنه كثيرا ما تصادفنا ظواهر ذات متغيرين أو أكثر، أو تتطلب دراستها وجود عدة متغيرات تتماشى معاً، تربطها علاقة واضحة، بحيث زيادة أحدها أو نقصانه تؤثر في الآخر بالزيادة أو النقصان، وبمعنى آخر تغير أحدها يؤدي إلى تغير الآخر إما إيجابياً أو سلبياً، وتوجد الكثير من الظواهر من هذا النوع، فعلى سبيل المثال زيادة الدخل المتاح للفرد، لابد أن تقابلها زيادة في مصاريفه الاستهلاكية، وانخفاض دخله لابد أن يؤثر سلباً أيضاً على مصاريفه الاستهلاكية، وزيادة كميات التساقط في موسم فلاحي ما، لابد أن يقابلها إنتاج وفير من الحبوب ما لم يتدخل عامل آخر، والعكس صحيح، في مثل هذه الظواهر توجد علاقة طردية بين متغيراتها، وهو ما يعبر عنه بالارتباط الطردي، كما يمكن أن تكون علاقة عكسية بين المتغيرات بحيث زيادة أحدها تؤدي إلى نقصان الآخر، فزيادة سعر مادة غذائية ما مثلاً، تؤدي إلى نقصان الكميات المستهلكة منها ما لم تكن ضرورية، وهذا ما يعبر عنه بالارتباط العكسي.

إن الهدف من هذا الفصل هو ترجمة العلاقة التي يمكن أن توجد بين متغيرين إقتصاديين أو أكثر، من خلال المعطيات الرقمية لهذه المتغيرات، إلى علاقة رياضية تحدد بنوع من الدقة طبيعة تلك العلاقة، هل هي علاقة طردية أو عكسية؟، وما هي درجة ارتباطها، قوية أم ضعيفة؟ وقبل ذلك ماهي العلاقة الدالية بينها، هل هي خطية أو ما يصطلح عليه بعلاقة الإنحدار الخطي والتي هي موضوع هذا الفصل؟ أو هي علاقة غير خطية؟

وذلك كله لأجل الاستفادة في معرفة درجة إستجابة أحد المتغيرات عند تغير الأخرى بقيم ما، وبمعنى آخر إستخدام العلاقة المستنتجة في التخمين المستقبلي للظواهر.

في علاقات الظواهر الفيزيائية، يمكن أن نصادف مثل هذه العلاقات، بحيث تكون في شكل خطي تام، كعلاقة المسافة المقطوعة بالنسبة للزمن، عند ثبات السرعة مثلاً، بحيث عند رسم هذه العلاقة على معلم متعامد تظهر نقاط الأزواج المرتبة على إستقامة تامة، وتكون دالة المسافة بالنسبة للزمن دالة خطية تامة، وهذا ما سنصطلح على تسميته بالإنحدار الخطي التام، غير أن مثل ذلك نادر المصادفة في علاقات المتغيرات الإقتصادية والاجتماعية عامة، إذ العلاقة بين المتغيرات في مثل هذه الظواهر يمكن أن تأخذ اتجاهها خطياً لكن ليس بالتام، ويتم إيجاد دالة الإنحدار التقريبية فقط.

ولمعرفة ما إذا كان الإنحدار الخطي تاماً أو غير تام، فإنه يتم أولاً تحديد طبيعة المتغيرات، أي ماهو المتغير الذي يؤثر في الآخر، وبمعنى آخر ما هو المتغير التابع و ماهو المتغير أو المتغيرات المستقلة أي المتغير أو المتغيرات التي تؤثر في التابع، ثم يتم على معلم متعامد تحديد نقاط العلاقة بين المتغيرات وهو ما نصلح عليه بشكل الانتشار، ومنه يمكن معرفة طبيعة الإنحدار، هل هو إنحدار خطي أو غير خطي، وهل هو إنحدار تام أو إنحدار غير تام.

نستشف من هذا التقديم أن هناك نوعين أساسيين من الإنحدار، الأول نصلح عليه بالإنحدار الخطي البسيط والذي يعتمد على متغيرين فقط أحدهما تابع و الآخر مستقل، و الثاني هو ما سنصلح عليه بالإنحدار المتعدد و الذي يعتمد على متغير تابع لمجموعة من المتغيرات المستقلة الأخرى، قد يكون عددها

إثنان أو ثلاثة أو أكثر من ذلك، ونظراً لأن دراسته تتطلب الإعتناء على جبر المصفوفات، فإننا نتطرق في هذا الفصل فقط إلى ما نسميه بالإنحدار الثلاثي وهو الذي يعتمد على متغير تابع ومتغيرين مستقلين إضافة إلى الإنحدار الخطي البسيط، وذلك في البند الأول، أما في البند الثاني فسوف نتطرق إلى معرفة كيفية إكتشاف درجة قوة العلاقة بين المتغيرات وهو ما نسميه بالارتباط .

أولاً : الإنحدار: سوف نتطرق في هذا البند إلى الإنحدار الخطي البسيط والإنحدار الخطي الثلاثي، متجنبين بذلك الإنحدار المتعدد لكونه يعتمد على جبر المصفوفات، وهو ما لم نعهد له . نشير إلى أن دراسة الإنحدار بصفة عامة تعتمد على فرضيات أساسية، تخص حد الخطأ أو البواقي، غير أننا سوف لن نتطرق لها في هذا المقام لكونها تحتاج إلى تسيقات نظرية أساسها الإحصاء الرياضي.

1- الإنحدار الخطي البسيط

أ- المتغيرات: يعرف في الإنحدار الخطي البسيط متغيرين أساسيين، إذا كانت لدينا القيم : $y_1, y_2, y_3, \dots, y_n$ ، تنجم عن القيم : $x_1, x_2, x_3, \dots, x_n$ ، على التوالي، فإنه يمكن إعطاء التعريف التالي :

تعريف 1-7: يسمى المتغير بالمتغير التابع، ونرمز له بـ: y_i ، إذا كانت كل قيمة له تتأثر تبعاً لقيمة متغير آخر x_i يسمى بالمتغير المستقل.

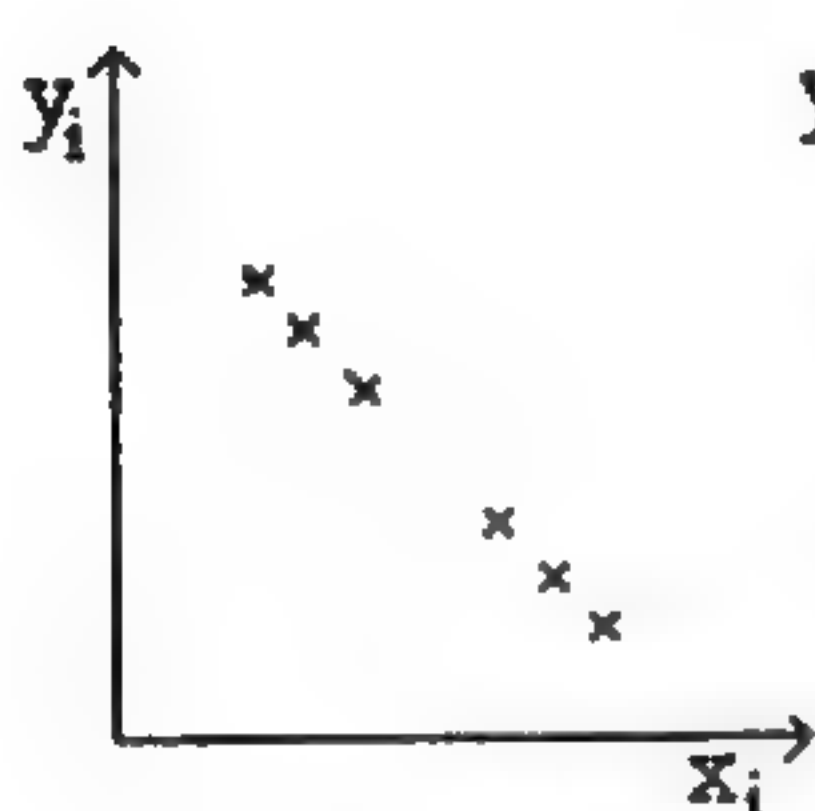
ب- شكل الإنحدار: بعد تحديد المتغير المستقل والمتغير التابع، يتم رسم معلم متعامد، بحيث يوضع على المحور العمودي المتغير التابع، وعلى المحور الأفقي المتغير المستقل، ويتم بعد ذلك تحديد نقاط الأزواج المرتبة

(x_i, y_i) ، لكل قيم الظاهرة . تسمى النقاط المحصل عليها، بشكل الانتشار .

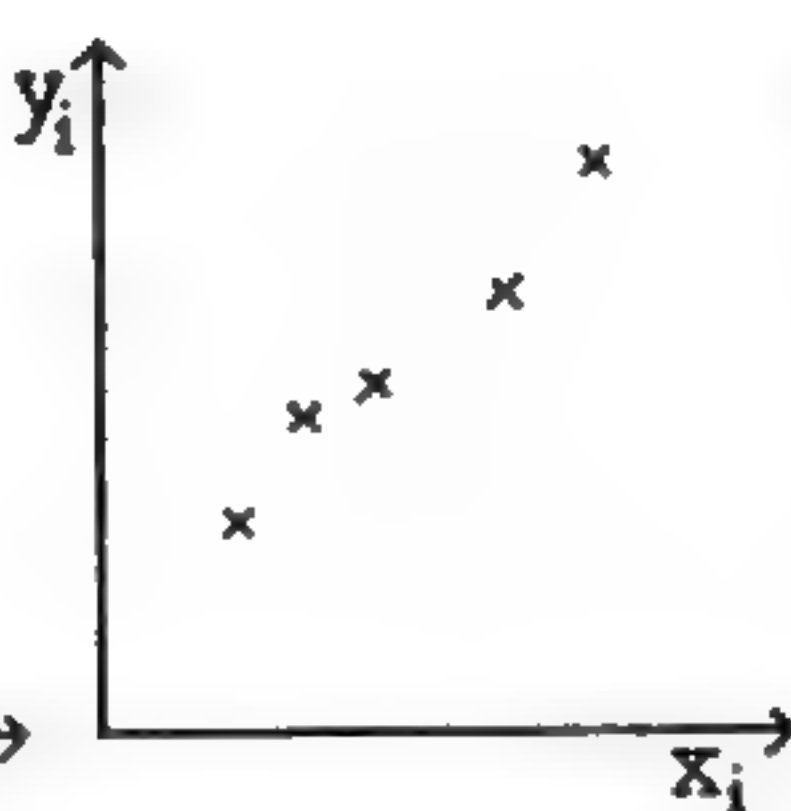
تعريف 2-7: شكل الانتشار لظاهرة ما ذات متغيرين أحدهما تابع و الآخر مستقل، هو مجموعة نقاط الأزواج المرتبة (x_i, y_i) ، لتلك الظاهرة المجسدة على معلم متعامد.

و يمكن أن نصادف عمليا في الظواهر ذات المتغيرين، عدة أنواع من أشكال الانتشار، كل نوع يحدد طبيعة الارتباط بينهما، وبالتالي يحدد طبيعة الإنحدار بينهما، ومن تلك الأشكال ما يلي:

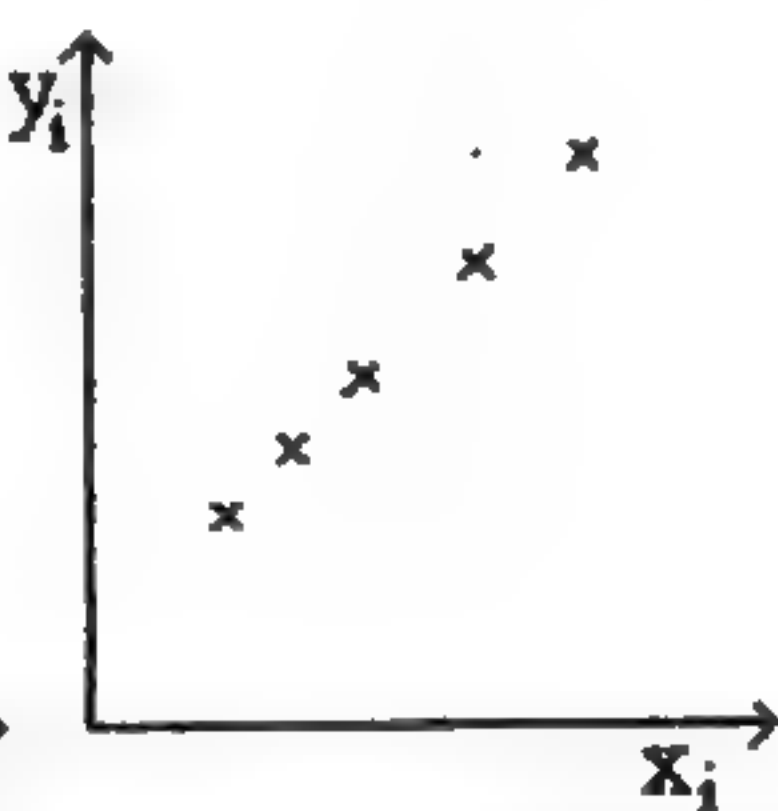
أمثلة 1-7



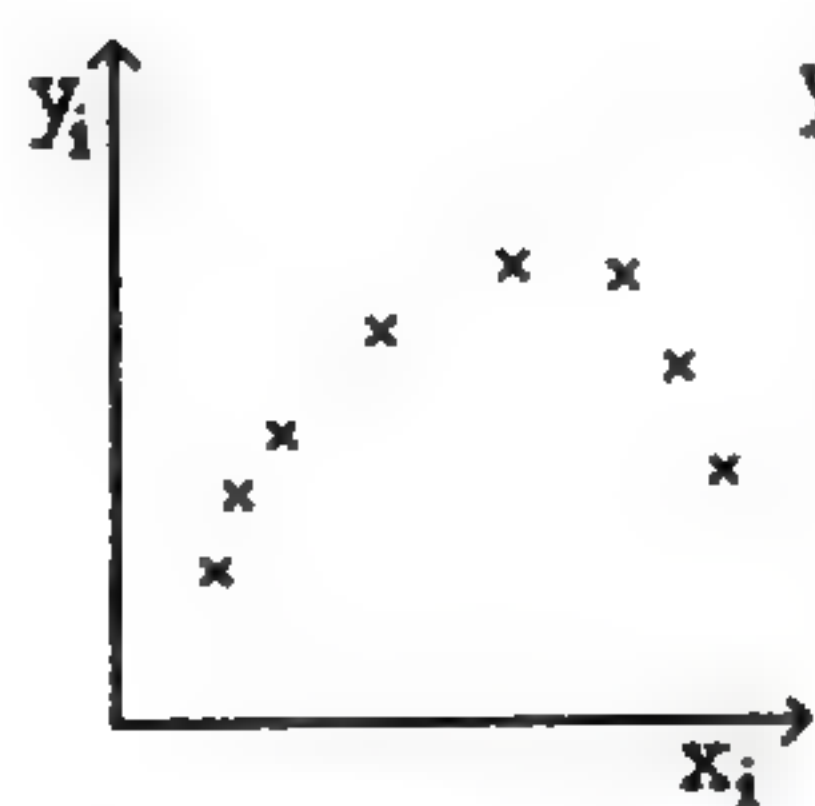
ارتباط خطي سالب تام
شكل 3-7



ارتباط خطي موجب غير تام
شكل 2-7



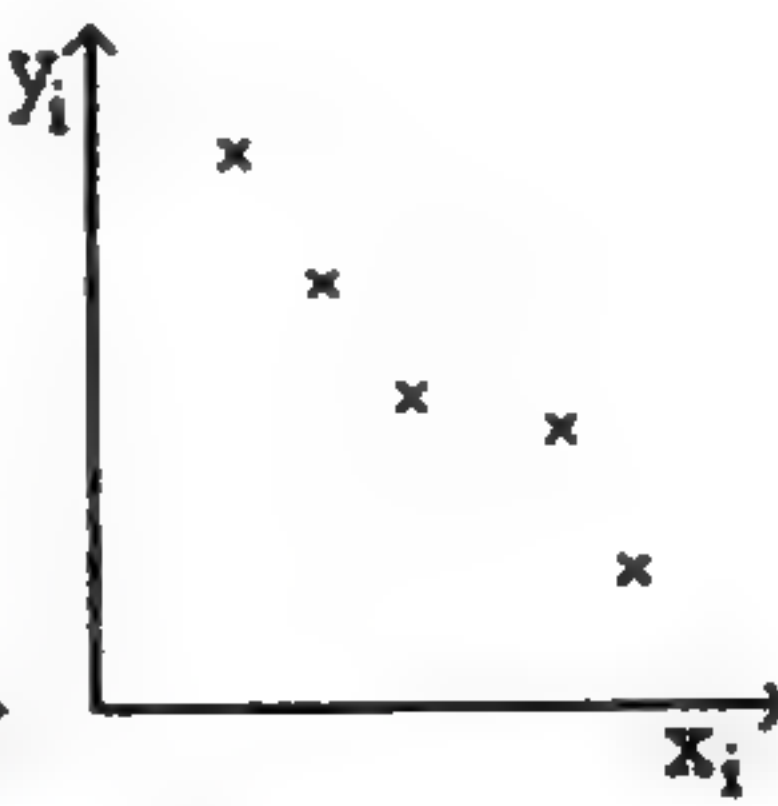
ارتباط خطي موجب تام
شكل 1-7



ارتباط غير خطي
شكل 6-7



لا يوجد ارتباط خطي
شكل 5-7



ارتباط خطي سالب غير تام
شكل 4-7

الشكلين 1-7 و 3-7، يوحيان بوجود إرتباط خطي تام بين المتغيرين، وبالتالي فإنه توجد علاقة بينهما تسمى علاقة إنحدار خطي تام، موجب (طردي) كما في الشكل الأول، وسالب (عكسي) كما في الشكل الثاني، وهذا النوع مسن الإنحدار قليل المصادفة في الظواهر الاقتصادية والاجتماعية، أما الشكلين 2-7 و 4-7 فيوحيان بوجود إرتباط غير تام بين المتغيرين، وبالتالي وجود علاقة إنحدار خطي غير تامة بينهما، طردية كما في الأول وعكسية كما في الثاني، وهذا النوع من العلاقات كثير المصادفة في الظواهر الاقتصادية والاجتماعية . وفي أحيان أخرى عند رسم شكل الانتشار نصادف شكل يشبه الشكل رقم 6-7 وفي هذه الحالة نقول أنه يوجد إرتباط بين المتغيرين غير أن هذا الارتباط يوحى بوجود علاقة إنحدار غير خطية بينهما بل وجود علاقة أسية، أما إذا كان شكل الانتشار مبعثرا كما في الشكل 5-7، فيدل ذلك على أنه لا توجد أية علاقة بين المتغيرتين . وتقاس شدة العلاقة بين المتغيرات المشار إليها بما يسمى بمعامل الارتباط، وهو رقم نسبي يوحى بطبيعة وشدة (درجة) العلاقة بينها (انظر البند الثاني).

ج- الإنحدار الخطي البسيط التام.

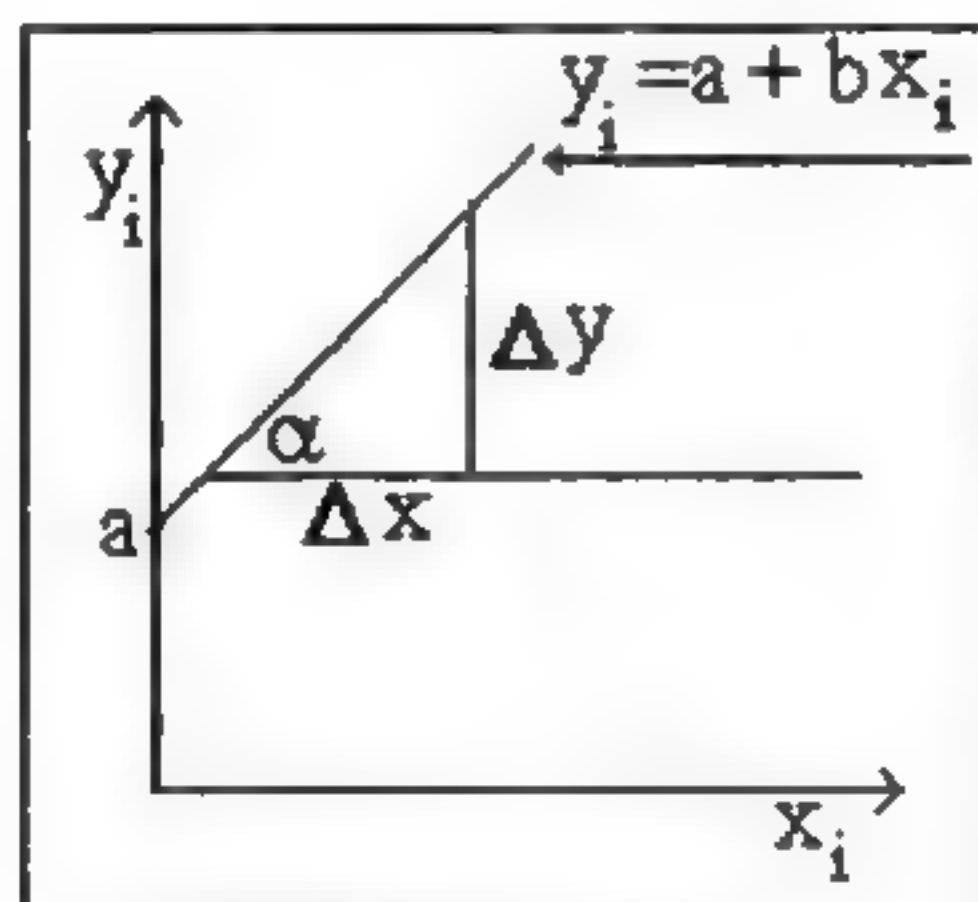
تعريفه 3-7: الإنحدار الخطي البسيط التام، هو الذي تكون كل نقاط شكل إنتشاره المحصل عليها من تجسيم متغيرتيه على معلم متعامد، على إستقامة تامة، سواء كانت في الاتجاه الموجب أو في الاتجاه السالب، وتكون معادلة الإنحدار على النحو التالي:

$$y_i = a + bx_i$$

1-7

حيث: b : ميل دالة الانحدار و a : ثابت الدالة، وتدل على قيمة الدالة عندما ينعدم x_i .

وهي دالة خط مستقيم يمكن إيجاد معالمها: b و a بسهولة، وذلك باستخدام الطرق الهندسية، إذ يتم الاتصال بين نقاط شكل الانتشار لنحصل على خط مستقيم ثم نوجد ميله، الذي يساوي إلى ظل الزاوية المحصورة بين المنحني والمستقيم الأفقي الموازي لمحور السينات، ونوجد بعد ذلك ثابت الدالة a ، وهو عبارة عن نقطة تقاطع المنحني مع المحور العمودي كما يوضحه الشكل الموالي.



شكل 7-7

بما أن b هو ميل الدالة، فهو إذن عبارة عن ظل الزاوية α المحصورة بين المنحني والمستقيم الموازي للمحور الأفقي، أي:

$$b = \frac{\Delta y}{\Delta x}$$

2-7

حيث أن القيم تكون موجودة على المعلم، لذلك يتم إيجاد b هندسيا بكل سهولة، بينما a هي عبارة عن نقطة تقاطع خط الانحدار مع المحور العمودي، ويمكن إيجادها هندسيا من الرسم أيضا، ونحصل بذلك على معلمي الدالة.

مثال 7-2: البيانات التالية تظهر تطور الدخل والاستهلاك لدولة ما بملايير الدينارات:

رقم السنة	الدخل	الإستهلاك
1	0	50
2	60	90
3	120	130
4	150	150
5	180	170
6	240	210
7	300	250

جدول 7-1

المطلوب:

- حدد المتغير التابع والمتغير المستقل.
- ارسم شكل الانتشار. ماذا تلاحظ ؟.
- أوجد معادلة إنحدار الاستهلاك على الدخل.

الاجابة:

أ- تحديد المتغير التابع و المتغير المستقل: بما أن الاستهلاك يتأثر بالدخل نظريا، إذ كلما ازداد الدخل ازداد الاستهلاك تبعا لذلك، والعكس صحيح و هو ما نلاحظه من خلال البيانات الإحصائية الواردة في الجدول، لذلك فإن:

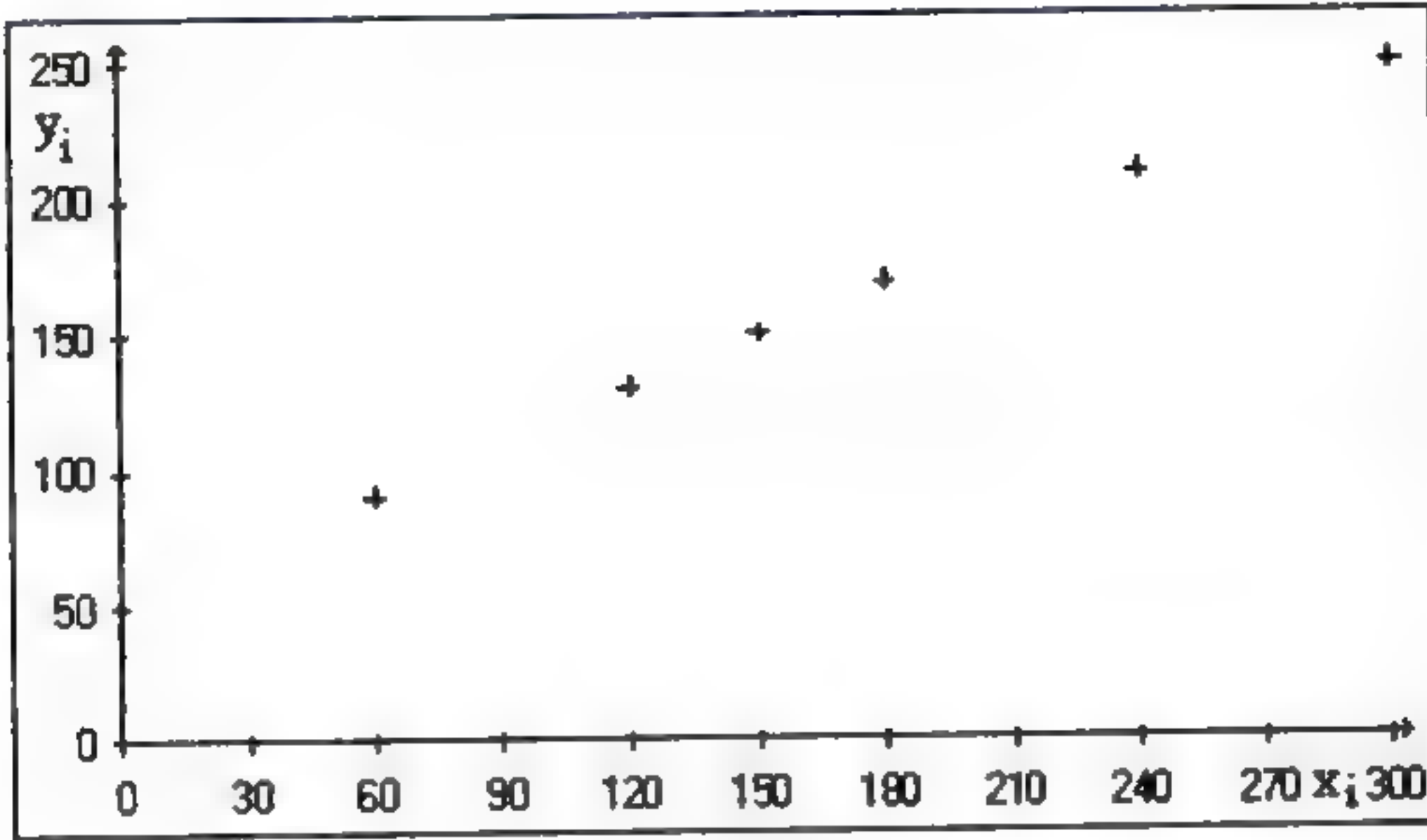
الإستهلاك : هو المتغير التابع ، ونرمز له بـ: y_i

الدخل : هو المتغير المستقل ، و نرمز له بـ: x_i

و يعني ذلك أن الإستهلاك هو دالة في الدخل أي:

$y_i = f(x_i)$ لا يمكننا أن نعرف طبيعتها إلا بعد أن نرسم شكل إنتشارها.

ب- شكل الانتشار:

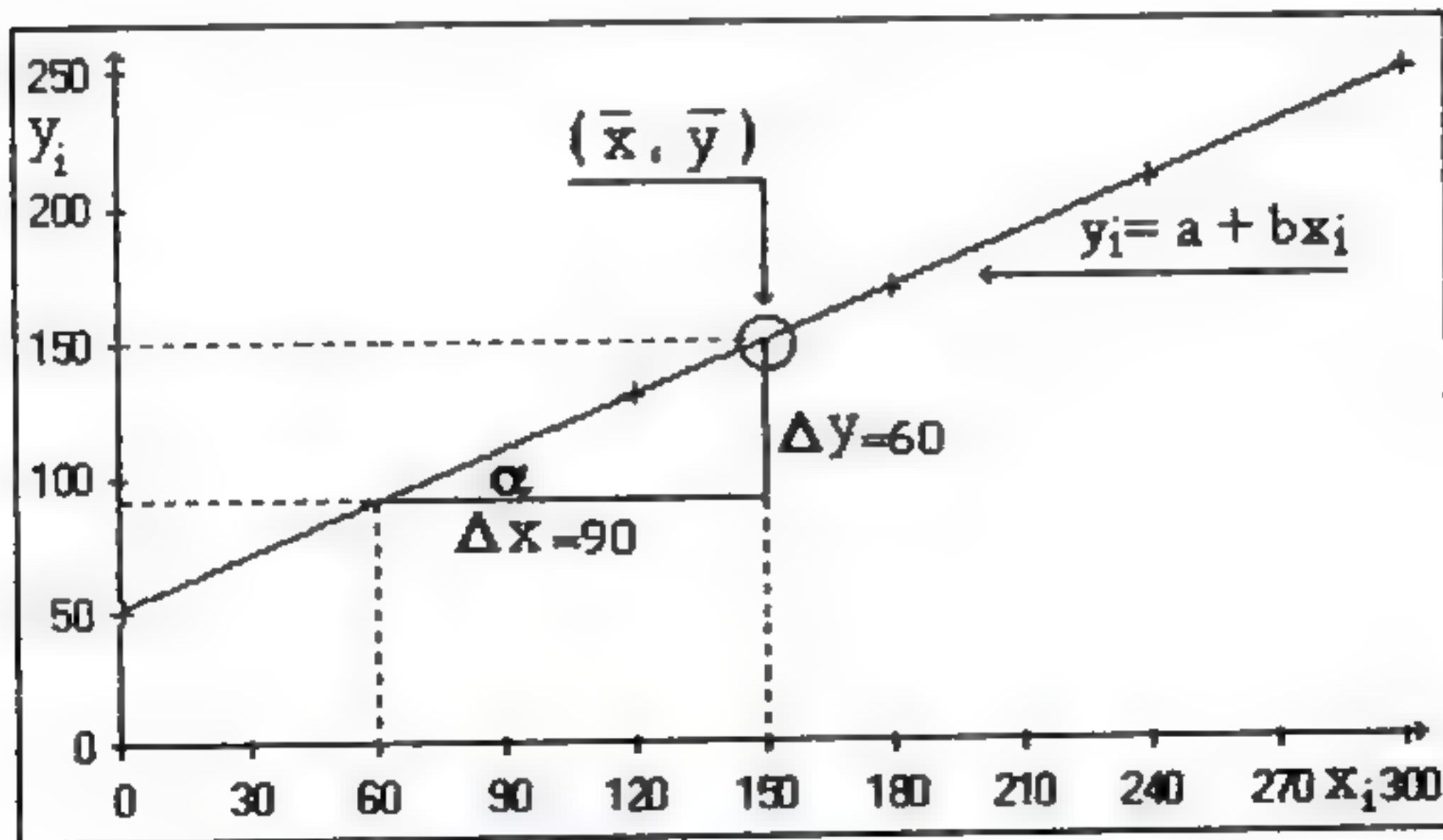


شكل 7-8

نلاحظ أن نقاط شكل الانتشار كلها على إستقامة تامة ، وبالتالي فإن معادلة الانحدار تكون على شكل المعادلة 1-7 ، أي:

$$y_i = a + bx_i$$

ج- إيجاد معادلة انحدار الاستهلاك على الدخل : بما أن نقاط شكل الانتشار على إستقامة تامة ، لذلك يتم إيجاد معلمي الدالة بالطرق الهندسية وبمساعدة الشكل 7-9 أدناه.



شكل 7-9

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018
حيث:

b هو ميل الدالة و ظل الزاوية α أي:

$$b = \text{Tga} = \frac{60}{90} = 0.67$$

a : هي نقطة تقاطع لمنحنى مع محور العينات و منه نجد: $a=50$

و بالتالي فإن معادلة إنحدار الإستهلاك على الدخل هي:

$$y_i = 50 + 0.67x_i \quad \text{مليار دينار} \quad 3-7$$

و يمكن التأكد من صحة المعادلة بإعطاء قيم لـ x_i من الجدول رقم 1-7، فسوف نجد بالضرورة نفس القيمة المقابلة لها.

و يمكن إستخدام هذه المعادلة في إيجاد أية قيمة مستقبلية للإستهلاك إذا عرف الدخل و ذلك بالتعويض في المعادلة، فإذا فرضنا أن الدخل سوف يصبح في السنة التاسعة 400 مليار دينار، فإن الإستهلاك المتوقع لهذه السنة هو:

$$y_9 = 50 + 0.67(400) = 318 \quad \text{مليار دينار}$$

يعني هذا أن الإستهلاك خلال السنة التاسعة سيكون 318 مليار دينار.

هذه هي الحالة الأولى من الانحدار الخطي ، و هي الحالة التي تكون فيها نقاط شكل الانتشار على إستقامة تامة ، حيث نجد معادلة الانحدار بدقة تامة ، غير أن هذه الحالة نادرة المصادفة ، في الظواهر الاقتصادية والاجتماعية ، كما سبقت الإشارة الى ذلك ، إذ في الغالب تكون نقاط شكل الانتشار ليست على إستقامة تامة ، لكنها تأخذ إتجاها عاما يقترب من الخط المستقيم ، وهو ما ستعرض له في البند الموالي .

د-الانحدار الخطي البسيط غير التام : في هذه الحالة

يشبه شكل الانتشار الشكل رقم 2-7 أو 4-7، إذ أن نقاط

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018
شكل الانتشار لا تكون على إستقامة تامة، لكنها تأخذ إتجاها
يمكن تقريبه من معادلة خط مستقيم، حيث أنه يستحيل إيجاد
المعادلة الحقيقية، التي نفترضها كما يلي:

$$y_i = a + bx_i + \varepsilon_i \quad 4-7$$

حيث أن الشق الأول من المعادلة وهو : $a + bx_i$ ، هو معادلة
خط مستقيم، وقد أضيف لها المقدار ε_i الذي هو قيمة البعد بين
النقاط الحقيقية ومعادلة الخط المستقيم، وحيث أنه يستحيل
إيجاد هذه المعادلة لذلك يتم تقريها تقديرا إلى المعادلة الخطية
التالية :

$$\hat{y}_i = \hat{a} + \hat{b}x_i \quad 5-7$$

حيث:

\hat{y} قيمة تقديرية لـ y_i و \hat{a} قيمة تقديرية لـ a و \hat{b} قيمة تقديرية لـ b
يمر مستقيم هذه الدالة من النقطة : \bar{x} و \bar{y} (الوسطين الحسابيين للمتغيرين).
يفترض أن تعطي هذه المعادلة خطا مستقيما أقرب ما يمكن إلى
جميع نقاط شكل الانتشار الحقيقية، و يتم إيجاد معلمتيها
(الثوابت)، بما يسمى بطريقة المربعات الصغرى، التي تعطي
المعادلتين التقديريتين لهما كما يلي:

$$\hat{b} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad 6-7$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x} \quad 7-7$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{N}$$

حيث: \bar{y} : الوسط الحسابي لقيم العينات، أي :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N}$$

: الوسط الحسابي لقيم السينات، أي :

و ستم البرهنة واشتقاق المعادلتين بطريقة المربعات الصغرى في البند الموالي.

مثال 7-3: البيانات التالية تظهر تطور كل من الدخل الداخلي والواردات السلعية بملايير الدينارات، خلال الفترة 2004/1997

السنة	واردات	د. داخلي
1997	10	10
1998	12	15
1999	15	20
2000	16	25
2001	16	30
2002	20	35
2003	26	40
2004	30	45

جدول 7-2

المطلوب:

- 1- حدد المتغير التابع والمتغير المستقل.
- 2- على معلم متعامد إرسم شكل الانتشار . ماذا تستنتج ؟
- 3- أوجد معادلة انحدار الواردات على الدخل الداخلي.

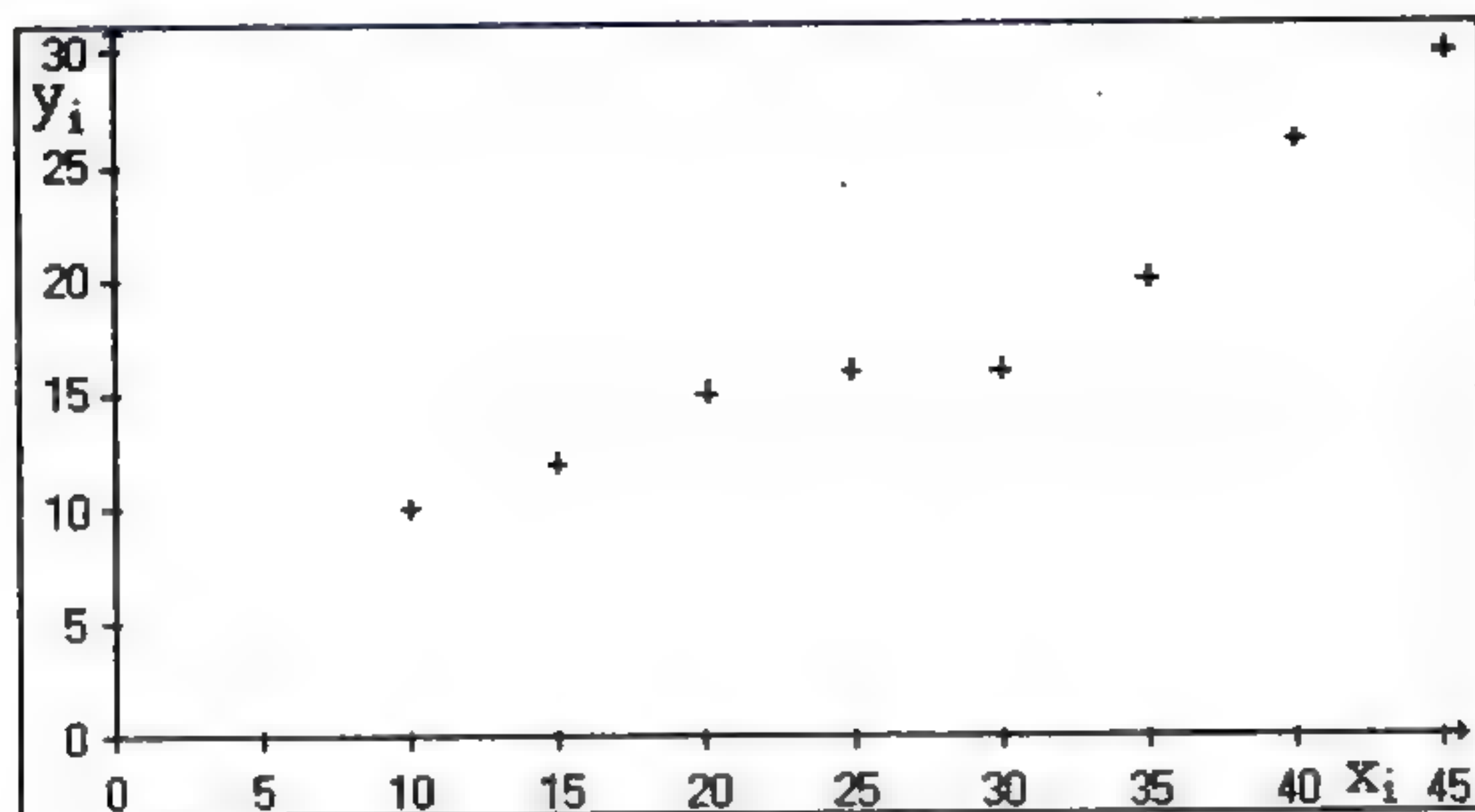
الإجابة :

*- الدخل الداخلي يعكس قدرة الدولة على تلبية الحاجيات، فكلما ازداد الدخل الداخلي تزداد قدرة الدولة على الاستيراد، و كلما إنخفض الدخل الداخلي إنخفضت قدرة الدولة على الاستيراد، و عليه فإن الواردات تزداد كلما ازداد الدخل الداخلي، وتنخفض كلما انخفض نظرياً، لذلك فإن :

- الواردات : هي المتغير التابع، ونرمز له بـ: y_i

- الدخل الداخلي : هو المتغير المستقل، ونرمز له بالرمز: x_i .

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018
* - شكل الانتشار:



شكل 7-9

نلاحظ أن نقاط شكل الانتشار ليست على استقامة تامة، غير أنه يمكن تقريبها الى معادلة خط مستقيم من شكل المعادلة رقم 7-5، وسيتم تقدير معالمها عن طريق المعادلتين : 7-6 و 7-7 المشار اليهما أعلاه، وبمساعدة الجدول التالي:

i	y_i	x_i	$(x_i - \bar{x})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
1	10	10	306.25	142.28
2	12	15	156.25	76.63
3	15	20	56.25	23.48
4	16	25	6.25	5.33
5	16	30	6.25	-5.33
6	20	35	56.25	12.53
7	26	40	156.25	98.38
8	30	45	306.25	207.73
مج	145	220	1050.00	561.60

جدول 7-3

نجد : الوسطين الحسابيين للمتغيرين المستقل و التابع هما:

$$\bar{x} = \frac{\sum_{i=1}^8 x_i}{8} = \frac{220}{8} = 27.5 \quad \text{و} \quad \bar{y} = \frac{\sum_{i=1}^8 y_i}{8} = \frac{145}{8} = 18.13$$

ميل دالة الانحدار المقدرة (المعلمة : \hat{b}) نجده بتطبيق المعادلة 6-7 أعلاه ومن

$$\hat{b} = \frac{561.69}{1050} = 0.53 \quad \text{نتائج الجدول 3-7 نجد :}$$

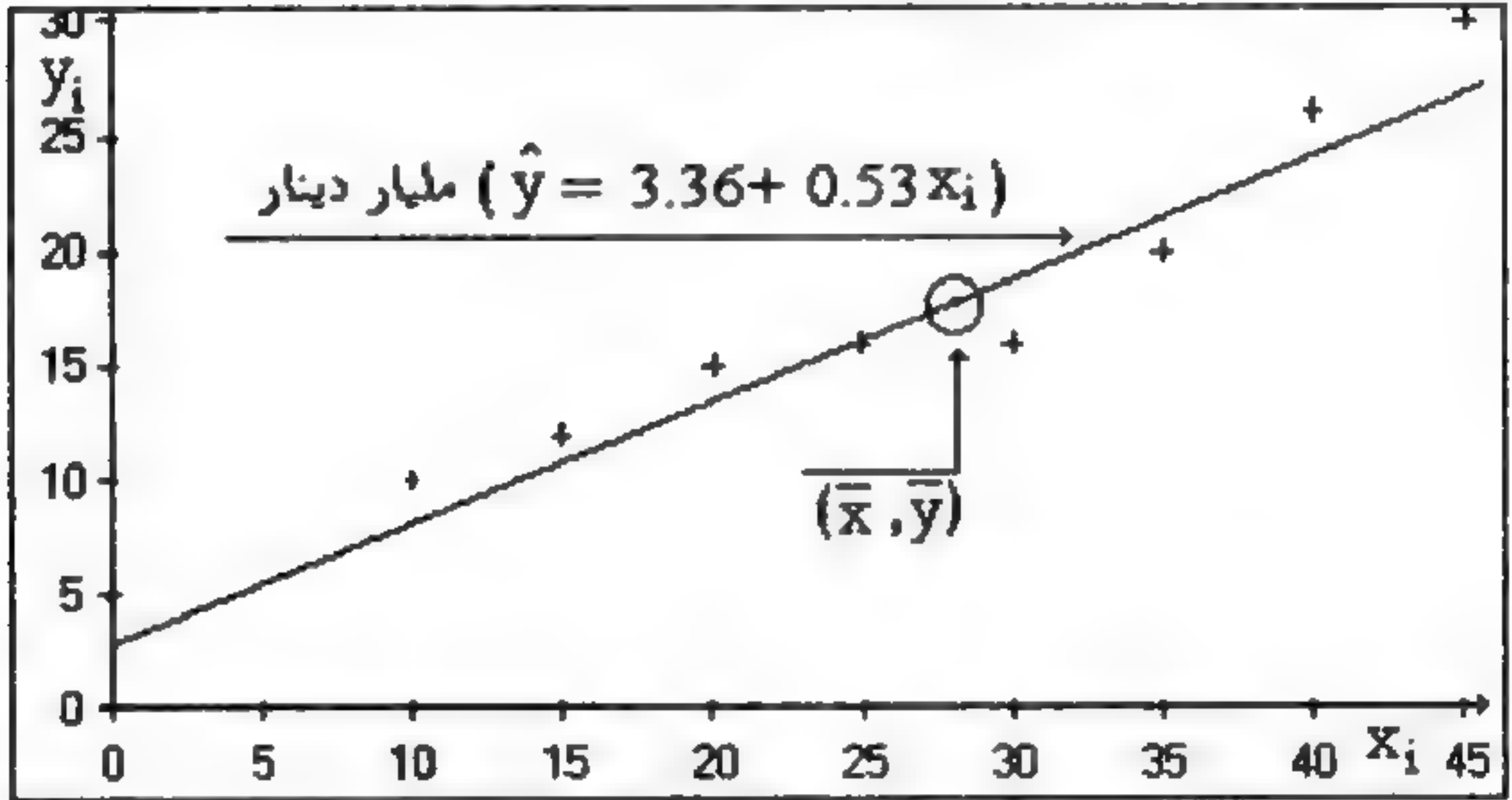
القيمة الثابتة لدالة الانحدار المقدرة (المعلمة : \hat{a}) هي :

$$\hat{a} = \bar{y} - \hat{b}\bar{x} = 18.13 - 0.53(27.5) = 3.36$$

ومنه تكون معادلة الانحدار المقدرة كما يلي:

$$(\hat{y}_i = 3.36 + 0.53x_i) \quad \text{مليار دينار}$$

يمكن ملاحظة أن مستقيم هذه المعادلة هو أقرب خط مستقيم الى كل نقاط شكل الانتشار من خلال البيان التالي :



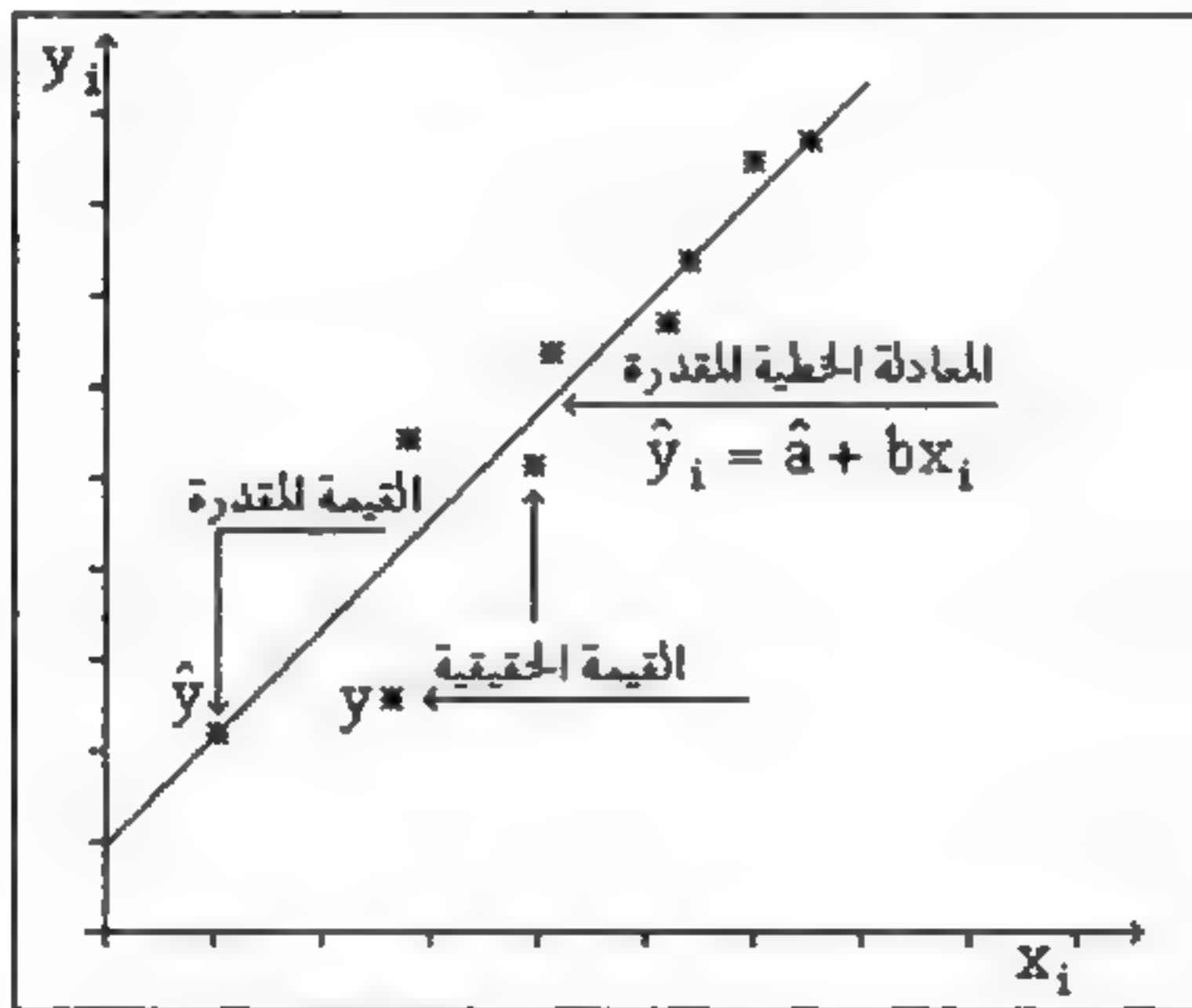
شكل 7-11

ويمكن أن يعطي قيمة تنبؤية للمتغير y ، عند اعطاء أية قيمة مستقبلية للمتغير x ، فإذا ما افترضنا على سبيل المثال أن قيمة الدخل الداخلي، ستكون 100 مليار دينار في سنة 2006، فإن القيمة التنبؤية للواردات في هذه السنة ستكون:

$$\hat{y}_{2006} = 0.53(100) + 3.36 = 56.36 \quad \text{مليار دينار}$$

و هذا ما يسمى بالتنبؤ النقطي، غير أن واقع الظواهر الاقتصادية و الإجتماعية يجعل من ذلك نادر المصادفة، لذلك ففي حالة التنبؤ يلجأ الى إستخدام ما يسمى بالجمال، و فكرته أن نوجد مجال تنبؤ أن تكون قيمة المتغير التابع ضمنه، و ذلك بإحتمال معين، و يتم ذلك أيضا باستخدام طرق علمية سوف لن نتعرض لها في هذا الكتاب.

تقدير المعامل بطريقة المربعات الصغرى: تهدف طريقة المربعات الصغرى كما سبقت الإشارة الى إيجاد معادلة خطية تقديرية، يكون خطها أقرب ما يمكن الى جميع نقاط شكل الانتشار على نحو الشكل التالي:



شكل 7-12

كما هو واضح في الرسم أعلاه، فإن معادلة المستقيم المفترض أن يكون أقرب ما يمكن الى جميع نقاط شكل الانتشار هي:

$$\hat{y}_i = \hat{a} + \hat{b} x_i$$

9-7

بينما المعادلة الحقيقية و التي يستحيل إيجادها ما دامت نقاط شكل الانتشار ليست على إستقامة تامة هي

$$y_i = a + bx_i + \varepsilon_i \quad 10-7$$

حيث ε_i هي قيمة عشوائية تعادل قيمة الانحرافات عن المنحنى الخطي للدالة. بإفتراض e_i هي قيمة تقريبية للمقدار ε_i و هي الفرق بين القيم الحقيقية y_i والقيم المقدرة : \hat{y}_i أي:

$$e = (y_i - \hat{y}_i) \quad 11-7$$

فإن هذا المقدار يسمى بالبواقي، و حتى تكون المعادلة التقديرية أقرب ما يمكن الى جميع نقاط شكل الانتشار، يجب أن يكون :

$$\sum_{i=1}^n e_i = \sum_{i=1}^n (y_i - \hat{y}_i) \rightarrow 0 \quad 12-7$$

أي مجموع البواقي يؤول الى الصفر، وهذا يكفي، كذلك مربعات البواقي تؤول الى أدنى قيمة ممكنة، أي:

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow \text{أدنى قيمة} \quad 13-7$$

بتعويض \hat{y}_i المقدرة بقيمتها حسب المعادلة 9-7 نجد :

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - a - \hat{b}x_i)^2 \rightarrow \text{أدنى قيمة} \quad 14-7$$

وحتى تأخذ المعادلة رقم 14-7 قيمتها الدنيا، يجب أن يتحقق شرطان :
الشرط الأول (اللازم): المشتقات الجزئية الأولى بالنسبة للمعالم، تساوي الصفر.

الشرط الثاني (الكافي): المشتقات الثانية بالنسبة للمعالم، يجب أن تكون أكبر من الصفر.

إذن لايجاد قيم المعالم المقدرة التي تجعل مربعات البواقي في أدنى قيمة لها، يجب أن تساوي المشتقة الأولى بالنسبة لميل الدالة الى الصفر، ثم نبحث عن قيمة الميل بدلالة بقية المتغيرات، ونقوم بنفس الشيء بالنسبة للمعلمة الثانية. نضع :

$$K = \sum_{i=1}^n (y_i - \hat{a} - \hat{b}x_i)^2 \quad 15-7$$

بالاشتقاق جزئيا بالنسبة لـ K :

$$\frac{\partial K}{\partial a} = -2 \sum_{i=1}^n (y_i - \hat{a} - \hat{b}x_i) = 0 \quad 16-7$$

ومنه يكون :

$$\sum_{i=1}^n (y_i - \hat{a} - \hat{b}x_i) = 0 \quad 17-7$$

$$\sum_{i=1}^n y_i - N\hat{a} - \hat{b} \sum_{i=1}^n x_i = 0 \quad 18-7 \text{ بفك القوس نجد}$$

بقسمة طرفي المعادلة 18-7 على المقدار N نجد:

$$\bar{y} - \hat{a} - \hat{b}\bar{x} = 0 \quad 19-7$$

و منه يكون :

$$\hat{a} = \bar{y} - \hat{b}\bar{x} \quad 20-7$$

باشتقاق K جزئيا بالنسبة للمعلمة الثانية نجد:

$$\frac{\partial K}{\partial b} = -2 \sum_{i=1}^n (y_i - \hat{a} - \hat{b}x_i)x_i = 0 \quad 21-7$$

$$\sum_{i=1}^n (y_i - \hat{a} - \hat{b}x_i)x_i = 0 \quad 22-7 \text{ أي}$$

بفك المعادلة 22-7 نجد :

$$\sum_{i=1}^n y_i x_i - \hat{a} \sum_{i=1}^n x_i - \hat{b} \sum_{i=1}^n x_i^2 = 0 \quad 23-7$$

وهذا يكافئ :

$$\sum_{i=1}^n y_i x_i = \hat{b} \sum_{i=1}^n x_i^2 + \hat{a} \sum_{i=1}^n x_i \quad 24-7$$

بتعويض المعادلة 20-7 في المعادلة 24-7 نجد:

$$\sum_{i=1}^n y_i x_i = \hat{b} \sum_{i=1}^n x_i^2 + (\bar{y} - \hat{b} \bar{x}) \sum_{i=1}^n x_i \quad 25-7$$

ومنه نجد:

$$\sum_{i=1}^n y_i x_i = \hat{b} \sum_{i=1}^n x_i^2 + \bar{y} \sum_{i=1}^n x_i - \hat{b} \bar{x} \sum_{i=1}^n x_i \quad 26-7$$

بإخراج \hat{b} عامل مشترك و تحويل $\bar{y} \sum_{i=1}^n x_i$ الى الطرف الأيسر نجد:

$$\sum_{i=1}^n y_i x_i - \bar{y} \sum_{i=1}^n x_i = \hat{b} \left(\sum_{i=1}^n x_i^2 - \bar{x} \sum_{i=1}^n x_i \right) \quad 27-7$$

وأخيرا تكون قيمة \hat{b} كما يلي :

$$\hat{b} = \frac{\sum_{i=1}^n x_i y_i - \bar{y} \sum_{i=1}^n x_i}{\sum_{i=1}^n x_i^2 - \bar{x} \sum_{i=1}^n x_i} \quad 28-7$$

يمكن إستخدام هذه المعادلة في تقدير ميل الدالة، غير أنه يصعب إستخدامها عندما تكون أرقام المتغيرات كبيرة، لذلك يتم إستخدام معادلة أخرى، تعتمد على الانحرافات عن الوسط الحسابي، وهي المعادلة المعطاة في البند السابق تحت رقم : 6-7 ، والتي يتم اشتقاقها باتباع المنهجية التالية:

في بسط المعادلة 28-7 نطرح ونضيف المقدار : $\bar{x} \sum_{i=1}^n y_i$

في مقام المعادلة 28-7 نطرح ونضيف المقدار : $\bar{x} \sum_{i=1}^n x_i$

فنحصل على العبارة التالية:

$$\hat{b} = \frac{\sum_{i=1}^n x_i y_i - \bar{y} \sum_{i=1}^n x_i - \bar{x} \sum_{i=1}^n y_i + \bar{x} \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - \bar{x} \sum_{i=1}^n x_i - \bar{x} \sum_{i=1}^n x_i + \bar{x} \sum_{i=1}^n x_i} \quad 29-7$$

الطرف الأخير في البسط هو : $N\bar{x}\bar{y}$ الطرف الأخير في المقام هو : $N\bar{x}^2$
 بالتعويض في 7-29، وإخراج المجموع نجد:

$$\hat{b} = \frac{\sum_{i=1}^n (x_i y_i - \bar{y} x_i - \bar{x} y_i + \bar{x} \bar{y})}{\sum_{i=1}^n (x_i^2 - 2\bar{x} x_i + \bar{x}^2)} \quad 30-7$$

ومنه يمكن كتابة البسط والمقام على النحو التالي :

$$\hat{b} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad 31-7$$

واضح أن المعادلة 7-31، سهلة الاستخدام، كما أنها تختصر الأرقام، لكونها تعتمد على الانحرافات عن الوسط الحسابي .
 تسمى طريقة إيجاد المعالم المقدرة لدالة الانحدار بهذه الطريقة، بطريقة المربعات الصغرى.

يمكن التأكد بأن القيمتين : \hat{a} و \hat{b} ، المقدرتين بهذه الطريقة تؤديان الى نهاية صغرى لمربعات البواقي، بمراجعة المشتقات الجزئية الثانية للمعادلة 7-15، وذلك بالنسبة للمعلمتين، اذ نجد هذه المشتقات موجبة.

كما هو ملاحظ فان الانحدار الخطي البسيط كما تم تناوله لحد الآن يعتمد على متغيرين أساسيين هما المتغير التابع و المتغير المستقل، غير أن هناك بعض المتغيرات الاقتصادية تعتمد على عدد أكثر من المتغيرات المستقلة، قد يكون عددها إثنان أو أكثر، ويتم تناول ذلك عادة في ما يسمى بالانحدار المتعدد، والذي قد يكون ثلاثيا أو رباعيا أو أكثر من ذلك، وذلك بالإعتماد على جبر المصفوفات، و نظرا لعدم تقديم تذكير في المصفوفات ضمن هذا الكتاب فإننا سوف نكتفي

2- **الإنحدار الثلاثي** : كما سبقت الإشارة أعلاه فإن الكثير من المتغيرات الاقتصادية تكون تابعة لمتغيرين مستقلين، وتكون علاقة الإنحدار على النحو التالي:

$$y_i = a + bx_i + cz_i + \varepsilon_i \quad 32-7$$

حيث : y_i : المتغير التابع.

x_i : المتغير المستقل الأول.

z_i : المتغير المستقل الثاني.

ε_i : حد الخطأ.

a, b, c : ثوابت.

و كما هو الشأن بالنسبة للإنحدار الثنائي إذ يستحيل إيجاد الدالة الحقيقية كما هي معرفة أعلاه، فإنه يلجأ الى تقديمها عن طريق دالة مقدرة كما يلي:

$$\hat{y}_i = \hat{a} + \hat{b}x_i + \hat{c}z_i \quad 33-7$$

و يتم إيجاد المعالم : \hat{a} ، \hat{b} و \hat{c} باستخدام طريقة المربعات الصغرى، بنفس المنهجية المستخدمة في الإنحدار الثنائي، إذ يكون الهدف هو تصغير مجموع مربعات انحرافات القيم المقدرة \hat{y}_i عن القيم الحقيقية y_i ، ويكون ذلك بإيجاد المشتقات الجزئية لمجموع مربعات الانحرافات و مساواتها الى الصفر، كما جرى في الإنحدار الثنائي، إذ يتم إستنتاج المعادلات التالية :

$$\sum_{i=1}^n y_i - \hat{b} \sum_{i=1}^n x_i - \hat{c} \sum_{i=1}^n z_i - N\hat{a} = 0$$

$$\sum_{i=1}^n y_i x_i - \hat{b} \sum_{i=1}^n x_i^2 - \hat{c} \sum_{i=1}^n z_i x_i - \hat{a} \sum_{i=1}^n x_i = 0$$

$$\sum_{i=1}^n y_i z_i - \hat{b} \sum_{i=1}^n x_i z_i - \hat{c} \sum_{i=1}^n z_i^2 - \hat{a} \sum_{i=1}^n z_i = 0$$

و حيث أنه لدينا ثلاثة مجاهيل هي : \hat{a} ، \hat{b} و \hat{c} و ثلاث معادلات، فإنه يمكن إيجاد المجاهيل بإحدى الطرق العادية ويكون :

$$\hat{b} = \frac{\sum_{i=1}^n y_i x_i \sum_{i=1}^n z_i^2 - \sum_{i=1}^n y_i z_i \sum_{i=1}^n x_i z_i}{\sum_{i=1}^n x_i^2 \sum_{i=1}^n z_i^2 - (\sum_{i=1}^n x_i z_i)^2} \quad 34-7$$

$$\hat{c} = \frac{\sum_{i=1}^n y_i z_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i x_i \sum_{i=1}^n x_i z_i}{\sum_{i=1}^n x_i^2 \sum_{i=1}^n z_i^2 - (\sum_{i=1}^n x_i z_i)^2} \quad 35-7$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x} - \hat{c}\bar{z} \quad 36-7$$

تقيس المعلمة \hat{b} التغير الحاصل في y بالنسبة لتغير مقدار x بوحدة واحدة عند ثبات z ، وبالمثل تقيس المعلمة \hat{c} التغير الحاصل في y بالنسبة لتغير z بوحدة واحدة عند ثبات x ، وتسمى \hat{b} و \hat{c} بمعاملات الإتحدار الجزئية.

مثال 4-7 : البيانات التالية تظهر أسعار محلات تجارية حسب مساحتها وعمر بنائها.

العمر (سنة)	المساحة (م ²)	السعر (10 ⁶ دج)
10	40	80
20	80	80
5	60	120
5	20	40
10	45	50
15	90	90
20	100	90
5	110	140

جدول 4-7

المطلوب:

1- حدد المتغير التابع مع التوضيح.

2- أوجد معادلة إنحدار السعر على كل من مساحة المحل عمره.

3- بناء على معادلة الإنحدار المحصل عليها، ماهو السعر المتوقع لمحل تجاري مساحته 75 م² وعمره 30 سنة.

الإجابة :

1- سعر المحل يتحدد نظريا حسب مساحته و عمره، إذ يتوقع أن يكون السعر مرتفعا كلما كانت مساحته كبيرة، وينخفض كلما كانت مساحته صغيرة، فالعلاقة بين السعر والمساحة هي علاقة طردية، بينما تكون العلاقة بين السعر والعمر عكسية، إذ يتوقع أن يكون سعر المحل منخفضا كلما كبر عمره وأن يكون السعر مرتفعا كلما قل عمره، فإذا رمزنا للسعر بالحرف y وللمساحة بالحرف x و للعمر بالحرف z ، فإن معادلة إنحدار السعر على المساحة وعلى العمر تكون على النحو التالي :

$$\hat{y}_i = \hat{a} + \hat{b}x_i + \hat{c}z_i$$

2- يتم إيجاد معالم الدالة المشار إليها أعلاه أي إيجاد معادلة الإنحدار عن طريق المعادلات : 34-7، 35-7، 36-7 وبمساعدة الجدول التالي:

i	y_i	x_i	z_i	$y_i x_i$	$y_i z_i$	$x_i z_i$	x_i^2	z_i^2
1	80	40	10	3200	800	400	1600	100
2	80	80	20	6400	1600	1600	6400	400
3	120	60	5	7200	600	300	3600	25
4	40	20	5	800	200	100	400	25
5	50	45	10	2250	500	450	2025	100
6	90	90	15	8100	1350	1350	8100	225
7	90	100	20	9000	1800	2000	10000	400
8	140	110	5	15400	700	550	12100	25
مج	690	545	90	52350	7550	6750	44225	1300

جدول 5-7

نجد :

$$\hat{a} = 7$$

$$\hat{b} = 1.43$$

$$\hat{c} = -1.63$$

و بالتالي تكون دالة إنحدار السعر على كل من المساحة والعمر هي :

$$\hat{y}_i = (7 + 1.43x_i - 1.63z_i).10^6 \text{ دج}$$

يعني هذا أن سعر المحل يساوي الى 7 مضافا اليه جزءا متعلقا بالمساحة ومطروحاً منه جزء آخر متعلقا بالعمر، ويلاحظ أن إشارة معامل x_i موجبة لتدل على أن العلاقة طردية بين السعر والمساحة، بينما إشارة معامل z_i سالبة لتدل على العلاقة العكسية بين السعر والعمر.

3- السعر المتوقع لمحل تجاري مساحته 75 م² و عمره 30 سنة يتم إيجاده بالتعويض في دالة الإنحدار المحصل عليها و يكون :

$$\hat{y}_i = 7 + 1.43(75) - 1.63(30) = 65.35.10^6 \text{ دج}$$

أي أن سعر المحل سيكون 65.35 مليون دينار

ثانياً: معاملات الارتباط : كما سبقت الإشارة عند دراسة الإنحدار، فإن المتغيرات الاقتصادية والاجتماعية كثيراً ما تكون متأثرة ببعضها البعض، وبمعنى آخر مرتبطة ببعضها البعض غير أن هذا الارتباط قد يكون قويا أو ضعيفا كما قد يكون موجبا أو سالبا (انظر الأشكال الواردة في الأمثلة (1-7)، وتقاس درجة الارتباط هذه بما يسمى بمعامل الارتباط، وهو على أنواع حسب عدد المتغيرات وطبيعتها، وسنتطرق فيما يلي الى كل من معامل الارتباط الخطي البسيط و معامل الارتباط الجزئي و كذا معامل ارتباط الرتب.

1- معامل الارتباط الخطي البسيط: إن درجة شدة العلاقة بين متغيرين أحدهما تابع والآخر مستقل، تقاس بما يسمى بمعامل الارتباط الخطي.

تعريف 4-7: اذا كانت لدينا مجموعة الأزواج المرتبة: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ فان شدة ارتباط المتغيرين ببعضهما البعض، تسمى بمعامل الارتباط الخطي البسيط، و يرمز له بالحرف r ، ويعطى بالمعادلة التالية:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{N \cdot \sigma_x \cdot \sigma_y} \quad 37-7$$

حيث: σ_x : الانحراف المعياري لقيم المتغير المستقل، σ_y : الانحراف المعياري لقيم المتغير التابع (أنظر مقاييس التشتت)

و يتميز معامل الارتباط بمجموعة من الخواص منها ما يلي:

أ- تتراوح قيمته بين -1 و $+1$ ، أي: $-1 \leq r \leq +1$

إذا كان: $r = 1$ ، فإن ذلك يدل على أن العلاقة بين المتغيرين هي علاقة خطية طردية تامة (ارتباط خطي موجب تام).

إذا ما كان: $r = -1$ ، فإن ذلك يدل على أن العلاقة بين المتغيرين هي علاقة خطية عكسية تامة (ارتباط خطي سالب تام).

إذا كان: $r = 0$ ، فإن ذلك يدل على أنه لا توجد علاقة خطية بين المتغيرين.

و عموماً نقول أن الارتباط بين المتغيرين هو ارتباط قوي، اذا كان:

$r \geq 0.7$ في حالة الارتباط الطردي.

$r \leq -0.7$ في حالة الارتباط العكسي.

ب- إذا كان \hat{b} معامل إنحدار y_i على x_i و \hat{b}^* معامل إنحدار x_i على y_i ، حيث:

$$\hat{b}^* = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad \hat{b} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$r^2 = \hat{b} \cdot \hat{b}^*$$

فإن:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad \text{و بالتالي نجد: 38-7}$$

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{b \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad \text{أو: 39-7}$$

و يمكن إستنتاج أيضا:

$$r = \hat{b} \frac{\sigma_x}{\sigma_y} \quad 40-7$$

حيث : σ_x : الانحراف المعياري لقيم المتغير المستقل، σ_y : الانحراف المعياري لقيم المتغير التابع.

ج- إذا كان : $r = 1$ فإن مجموع مربعات فروقات القيم الحقيقية عن القيم المقدرة يكون معدوما أي :

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = 0$$

يعني ذلك أن هناك علاقة ارتباط خطي تام بين المتغير التابع و المتغير المستقل.
و إذا كان : $r = 0$ فإن مجموع مربعات فروقات القيم الحقيقية عن القيم المقدرة تأخذ أكبر قيمة لها، ويعني ذلك أن المتغيرين مستقلين عن بعضهما البعض أو لا توجد علاقة إنحدار خطي بينهما.

تجدر الإشارة أخيرا الى أن : r تأخذ نفس اشارة معامل دالة الإنحدار.

مثال 5-7: أوجد معامل ارتباط بيانات المثال 3-7.

باستخدام المعادلة رقم 7-38 و بمساعدة الجدول التالي نجد:

i	y_i	x_i	$(x_i - \bar{x})^2$	$(x_i - \bar{x})(y_i - \bar{y})$	$(y_i - \bar{y})^2$
1	10	10	306.25	142.28	66.10
2	12	15	156.25	76.63	37.58
3	15	20	56.25	23.48	6.80
4	16	25	6.25	5.33	4.45
5	16	30	6.25	-5.33	4.45
6	20	35	56.25	12.53	2.79
7	26	40	156.25	98.38	61.94
8	30	45	306.25	207.73	140.90
مج	145	220	1050.00	561.69	328.37

جدول 6-7

$$r = \frac{561.69}{\sqrt{1050} \cdot \sqrt{328.37}} = 0.96$$

يعني هذا أن الارتباط بين المتغيرين الواردات والدخل الداخلي هو ارتباط طردي قوي، كما تعني القيمة :

$$r = 0.96 \text{ أو } r = 96\%$$

أن معادلة الإنحدار تفسر 96 % من التغير الإجمالي في المتغير المستقل أما النسبة المتبقية و هي 4% فإنها تعود الى عوامل متضمنة في حد الخطأ.

2- معامل الارتباط الجزئي : في الإنحدار الثلاثي يكون المتغير التابع دالة في متغيرين مستقلين، يؤثران بدرجة أو أخرى في المتغير التابع، ويكون من المفيد معرفة درجة الارتباط الصافي للمتغير التابع بكل متغير من المتغيرين المستقلين، ويتم ذلك عن طريق ما يسمى بمعامل الارتباط الجزئي.

تعريف 5-7 : معامل الارتباط الجزئي هو أداة لقياس صافي الارتباط بين متغير تابع، وآخر مستقل بعد حذف التأثير المشترك للمتغيرات المستقلة الأخرى، أي مع تثبيتها.

إذا كانت لدينا دالة الإنحدار المفترضة من الشكل :

بوابه

193 الباحث الاقتصادي

Economic Research Gate

$$y_i = a + bx_i + cz_i$$

فإن معاملات الارتباط الجزئي بين y_i و x_i و z_i على التوالي تعطى كما يلي:
أ- معامل الارتباط الجزئي بين y و x مع تثبيت z :

$$r_{yx.z} = \frac{r_{yx} - r_{yz} \cdot r_{xz}}{\sqrt{1 - r_{xz}^2} \cdot \sqrt{1 - r_{yz}^2}} \quad 41-7$$

ب- معامل الارتباط الجزئي بين y و z مع تثبيت x :

$$r_{yz.x} = \frac{r_{yz} - r_{yx} \cdot r_{xz}}{\sqrt{1 - r_{xz}^2} \cdot \sqrt{1 - r_{yx}^2}} \quad 42-7$$

حيث : r_{yx} : معامل الارتباط البسيط بين المتغيرين y و x .
 r_{yz} : معامل الارتباط البسيط بين المتغيرين y و z .
 r_{xz} : معامل الارتباط البسيط بين المتغيرين x و z .

كما هو الشأن بالنسبة للارتباط البسيط فإن معامل الارتباط الجزئي تتراوح قيمته أيضا بين -1 و $+1$ و تأخذ نفس إشارة المعلمة المقابلة في دالة الانحدار.
مثال 6-7: أوجد معاملات الارتباط الجزئي لبيانات المثال 4-7.

الإجابة :

أ- معامل الارتباط الجزئي بين y و x مع تثبيت z ، أي $r_{yx.z}$ يتم إيجاده باستخدام المعادلة رقم: 41-7، حيث نجد :

$$r_{yx} = 0.73 \quad r_{yz} = -0.14 \quad r_{xz} = 0.43$$

و بالتالي يكون : $r_{yx.z} = 0.88$

ب- معامل الارتباط الجزئي بين y و z مع تثبيت x ، أي $r_{yz.x}$ يتم إيجاده باستخدام المعادلة رقم: 42-7، حيث نجد :

$$r_{yz.x} = -0.74$$

تعني العبارة : $r_{yx.z} = 0.88$ بعد أن يكون عمر الخجل قد فسر ما يستطيع من التغير في السعر، فإن المساحة تفسر حوالي 88 % من التغير المتبقي في السعر. و بالمثل يفسر $r_{yz.x}$ ، غير أن العلاقة طردية بين السعر و المساحة تكون إشارة معامل الارتباط الجزئي موجبة،

وعكسية بين السعر و العمر لكون إشارة معامل الارتباط سالبة.

3-معامل ارتباط الرتبة : في بعض الحالات تكون لدينا بيانات وصفية يمكن التمييز بينها بمعرفة رتبها، كأن تكون لدينا مجموعة من المواد يمكن للمستهلك أن يربتها حسب أهميتها بالنسبة له، و في مثل هذه الحالات لا يمكن إيجاد معامل الارتباط كما تم تقديمه آنفاً، لكن يتم إيجاد معامل سبيرمان للرتب و الذي يعطى عن طريق المعادلة التالية :

$$r = 1 - \frac{6 \sum_{i=1}^n d_i^2}{N(N^2 - 1)} \quad 43-7$$

حيث: d الفرق بين ترتيب x و y.

ونحيث يكون هذا الترتيب تنازلياً، أي أن أكبر قيمة تأخذ الرتبة 1، وأقل قيمة تأخذ الرتبة الأخيرة.

و يصلح استخدام معامل ارتباط الرتب خاصة إذا ما كان N يتراوح بين 25 و 30، وتكون قيمته أيضاً محصورة بين: 1+ و 1-.

مثال 7-7: أوجد معامل ارتباط الرتب للبيانات التالية :

x	25	20	30	45	10	17	23	21
y	70	47	73	40	45	50	48	63

جدول 7-7

الإجابة : يكون ترتيب القيم x و y على النحو التالي :

ترتيب x	3	6	2	1	8	7	4	5
ترتيب y	2	6	1	8	7	4	5	3

جدول 8-7

وتتم بقية الحسابات من خلال الجدول التالي :

ترتيب x	3	6	2	1	8	7	4	5	مح
ترتيب y	2	6	1	8	7	4	5	3	
d	1	0	1	-7	1	3	-1	2	
d ²	1	0	1	49	1	9	1	4	66

جدول 9-7

$$r = 1 - \frac{6(66)}{8.(64 - 1)} = 0.21$$

بتطبيق المعادلة أعلاه نجد :

و هو ارتباط ضعيف.

تستخدم هذه الطريقة في حالة ما إذا كانت البيانات غير متساوية، أما إذا كان البعض منها متساو، فإننا نتبع الخطوات التالية :

1- نرتب القيم ونعطي لكل قيمة ترتيباً، كما لو أن ليس فيها قيم متساوية، بحيث أكبر قيمة تأخذ الرتبة 1 والقيمة الأقل تأخذ الرتبة الأخيرة.

2- نوجد الوسط الحسابي لرتب كل مجموعة من البيانات المتساوية، ونعطي لكل قيمة من القيم المتساوية ترتيباً يساوي هذا الوسط.

مثال 7-8: أوجد معامل ارتباط الرتب للبيانات التالية :

x	25	20	30	45	20	17	20	25
y	70	45	73	40	45	50	45	63

جدول 7-10

نلاحظ أنه توجد قيم متساوية في كل من المتغيرين x و y، لذلك نقوم بإيجاد متوسط الرتب للقيم المتساوية بعد ترتيبها، ونعطي كل قيمة من القيم المتساوية رتبة تساوي هذا المتوسط و ذلك عند موقعها الأصلي وذلك كما هو واضح في الجدول 7-11.

بالنسبة للمتغير x مثلاً، نجد أن القيمة 20 مكررة 3 مرات ترتيبها على التوالي هو : 5، 6، 7 لذلك يكون متوسط هذه الرتب هو : $6 = 3 \setminus (7+6+5)$ لذلك نعطي كل قيمة تساوي 20 عند موقعها في الجدول الأصلي ترتيباً يساوي 6، ونقوم بنفس الطريقة بالنسبة للقيم المتساوية الأخرى.

ترتيب x	3.5	6	2	1	6	8	6	3.5	مج
ترتيب y	2	6	1	8	6	4	6	3	
d _i	1.5	0	1	-7	0	4	0	0.5	
d ² _i	2.25	0	1	49	0	16	0	0.25	68.5

جدول 7-11

$$r = 1 - \frac{6(68.5)}{8.(64 - 1)} = 0.18$$

و بتطبيق المعادلة رقم 7-43 نجد :

تمارين.

تمرين 1: البيانات التالية تظهر تطور أسعار مادة ما بالدينار و الكميات المطلوبة منها في إحدى المدن بمئات الأطنان :

i	السعر	الكمية
1	80	130
2	85	130
3	87	125
4	100	120
5	110	90
6	123	80
7	140	72
8	135	70

- المطلوب:** 1- حدد المتغير التابع والمتغير المستقل.
- 2- على معلم متعامد ارسـم شكل الانتشار. ماذا تستنتج؟
- 3- أوجد معادلة انحدار الكميات على الأسعار، وفسرها.
- 4- أوجد الانحراف المعياري للأسعار وللكميات وفسره.
- 5- أوجد معامل الارتباط الخطي وفسره.
- 6- ماهي الكميات المتوقعة استهلاكها عند الأسعار : 180 و 50 دينار على التوالي.

تمرين 2: البيانات التالية تظهر تطور عدد السكان و عدد المواليد الأحياء (الزيادة الطبيعية) في الجزائر بين سنتي 1990 و 2002 في الجزائر.

السنة	1990	1991	1992	1993	1994	1995	1996
عدد السكان ¹⁰ نسمة	25.0	25.6	26.3	26.9	27.4	28.0	28.6
المواليد الأحياء ¹⁰	624	618	639	607	596	531	482
السنة	1997	1998	1999	2000	2001	2002	
عدد السكان ¹⁰ نسمة	29.0	29.5	30.0	30.4	30.8	31.2	
المواليد الأحياء ¹⁰	476	464	452	449	478	479	

- المطلوب:** 1- ما هو المتغير التابع و ماهو المتغير المستقل.
- 2- إرسـم شكل الإنتشار، ماذا تستنتج؟
- 3- أوجد معادلة إنحدار المواليد على عدد السكان.

4- ما هو عدد المواليد المتوقع إذا ارتفع عدد السكان الى 35 مليون نسمة في سنة 2007.

تمرين 3 : البيانات التالية تظهر تطور عدد السكان و عدد مناصب العمل المحققة في الجزائر خلال الفترة 1990-2001 حسب مصادر الديوان الوطني للإحصائيات.

السنة	1990	1991	1992	1993	1994	1995
عدد السكان 10 ⁶ نسمة	25.0	25.6	26.3	26.9	27.4	28.0
مناصب العمل	60498	42219	36668	35431	36985	41463
السنة	1996	1997	1998	1999	2000	2001
عدد السكان 10 ⁶ نسمة	28.6	29.0	29.5	30.0	30.4	30.8
مناصب العمل	32110	24830	26664	22377	22215	23696

المطلوب: 1- ما هو المتغير المستقل و المتغير التابع؟

2- إرسم شكل الانتشار. ماذا تستنتج؟

3- أوجد معادلة إنحدار مناصب العمل المنشأة على عدد السكان.

تمرين 4: البيانات التالية تظهر تطور كل من الإنتاج الداخلي الإجمالي والواردات، خلال الفترة 1976-1987 للجمهورية الجزائرية حسب أرقام الديوان الوطني للإحصائيات ، بملايير الدينارات.

السنة	الواردات	إنتاج د.إ.
1976	28.43	68.50
1977	37.27	81.90
1978	41.99	104.00
1979	46.47	128.50
1980	55.84	162.50
1981	65.99	191.50
السنة	الواردات	إنتاج د.إ.
1982	68.32	207.60
1983	66.72	239.80
1984	68.16	259.90
1985	65.06	289.20
1986	55.79	286.50
1987	48.88	307.90

- 1- الإنتاج الداخلي الإجمالي، والواردات السلعية متغيرتين إقتصاديتين، أيهما متغير تابع وأيهما متغير مستقل، حسب أصول النظرية الإقتصادية.
 - 2- قدم البيانات المجدولة في شكل بياني مناسب. ماذا تستنتج؟.
 - 3- إذا كانت الواردات السلعية دالة خطية في الناتج الداخلي الإجمالي، ماهي صيغتها العامة؟
 - 4- بناء على السؤال 3، أوجد الدالة المقدرة للواردات السلعية، وفسر معالمها المقدرة.
 - 5- أوجد درجة الارتباط بين المتغيرتين، وفسرها.
 - 6- إذا حددت قيمة الناتج الداخلي الإجمالي سنة 2006 بـ: 1200 مليون دج، أوجد الواردات المتوقعة خلال تلك السنة.
- تمرين 5:** البيانات التالية تظهر تطور الكميات المستهلكة من مادة القهوة Q_1 والكميات المستهلكة من مادة السكر Q_2 ، وسعر القنطار الواحد من مادة السكر p ، خلال الفترة 1998 \ 2005 في دولة ما.

السنة	Q_1 قنطار 10 ³	Q_2 قنطار 10 ³	p د ج/قنطار
1998	20	10	100
1999	30	17	107
2000	35	20	110
2001	40	30	120
2002	55	43	133
2003	50	50	140
2004	65	65	155
2005	60	75	165

- المطلوب: 1- الكميات المستهلكة من القهوة والسكر متغيرتين إقتصاديتين، أيهما متغير تابع وأيهما متغير مستقل. وضح.
- 2- على معلم متعامد أرسم شكل إنتشار كميات السكر على كميات القهوة. ماذا تستنتج؟

3- أوجد معادلة انحدار Q_2 على Q_1 وفسرها. ثم ارسمها على المعلم المطلوب في السؤال 2.

4- اشرح معامل الارتباط واحسبه لتغيرتي السؤال 1.

5- أوجد شكل انتشار الكميات المستهلكة من السكر على السعر. ماذا تستنتج؟

6- باستعمال الطريقة الهندسية أوجد انحدار الكميات المستهلكة من السكر على الثمن. فسرهما.

7- أوجد القيمة التنبؤية لإستهلاك السكر عندما يصبح السعر 200 دينار للقنطار.

8- إذا كان سعر السكر دالة في الكميات المستهلكة من المادتين أوجد الدالة المقدرة للسعر بدلالة كميات المادتين. 9- أوجد معاملات الارتباط الجزئي وفسرها.

تمرين 6: إذا كانت لدينا المعادلتين الطبيعيين:

$$\sum_{i=1}^n (y_i - \hat{a} - \hat{b}x_i)x_i = 0$$

$$\sum_{i=1}^n (y_i - \hat{a} - \hat{b}x_i) = 0$$

إثبت أنه يمكن كتابة \hat{a} و \hat{b} المقدرتين بطريقة المربعات الصغرى على نحو المعادلات التالية:

$$\hat{b} = \frac{\sum_{i=1}^n x_i y_i - y \sum_{i=1}^n x_i}{\sum_{i=1}^n x_i^2 - x \sum_{i=1}^n x_i}$$

$$\hat{b} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

تمرين 7 : إذا كانت معادلة الارتباط الخطي على النحو:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{N \cdot \sigma_x \cdot \sigma_y}$$

حيث : σ_x : الانحراف المعياري لقيم المتغير المستقل، σ_y : الانحراف المعياري لقيم المتغير التابع.

المطلوب: إثبت أنه يمكن كتابة معامل الارتباط أيضا على النحو:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$r = \sqrt{\hat{b} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

و على النحو:

$$r = b \frac{\sigma_x}{\sigma_y}$$

أو :

تمرين 8 :إليك البيانات التالية :

$$\sum_{i=1}^8 (y_i - \bar{y})^2 = 49748$$

$$\sum_{i=1}^8 (x_i - \bar{x})(y_i - \bar{y}) = 4127.5$$

$$\sum_{i=1}^8 y_i = 817$$

$$\sum_{i=1}^8 x_i = 860$$

$$r = 0.95$$

المطلوب: 1- أوجد دالة الإنحدار الخطي المقدرة بناء على هذه المعطيات.

3- أوجد الإنحراف المعياري لقيم المتغيرين التابع و المستقل.

الفصل الثامن

السلام الزمني

الكثير من الظواهر الاقتصادية والاجتماعية، تكون ذات تغيرات شهرية أو فصلية أو سنوية، نتيجة لتوافر مجموعة من الظروف في وقت معين، لو تتبعنا الإحصائيات المتعلقة بالكميات المستهلكة من الغاز مثلاً، لوجدنا أن الكميات المستهلكة منه ترتفع شتاء، ثم تبدأ في الانخفاض تدريجياً مع بداية فصل الربيع، لتصل إلى أدنى كمية مستهلكة صيفاً، ثم تبدأ الكميات المستهلكة في التزايد مع بداية فصل الخريف، وتكرر تقريراً نفس التواتر عبر فصول السنة، في مثل هذه الحالات تكون المتغيرة الاقتصادية مرتبطة بالفترة الزمنية، وتشكل مجموعة البيانات المحصل عليها لمثل هذه الظواهر ما يسمى بالسلسلة الزمنية.

تعريف 8-1: السلسلة الزمنية لظاهرة ما هي عبارة عن مجموعة من مشاهدات تلك الظاهرة مأخوذة خلال فترات زمنية متتالية وذات أبعاد متساوية، إذا كانت هذه المشاهدات هي :

للظاهرة هي : $y_1, y_2, y_3 \dots y_n$ مأخوذة خلال الفترات : $t = 1, 2, 3 \dots n$ فإن الدالة الزمنية

$$y_t = f(t) \quad 1-8$$

حيث : y_t : تسمى بالمتغير التابع . t : المتغير المستقل .
وقد تكون الفترات الزمنية مقاسة بالسنة أو أجزاءها،
أو أضعافها.

تهدف دراسة السلسلة الزمنية لظاهرة ما، الى تحديد كيفية تغير تلك الظاهرة عبر الزمن، والى تحديد دورات تلك التغيرات، ومعرفة أسبابها ونتائجها، وكذلك التخمين المستقبلي لتطورها.

عملياتا الكثير من شؤون التسيير في مختلف المؤسسات الإنتاجية والتجارية والاجتماعية، تستخدم السلاسل الزمنية سواء في توقعات الإنتاج أو المبيعات أو الاتجاه المستقبلي لظواهرها، الشيء الذي يسمح للقائمين على هذه المؤسسات باتخاذ التدابير اللازمة لمواجهة أي طارئ.

أولاً: أشكال تغيرات السلسلة: لتحديد طبيعة السلسلة الزمنية، يتم إعداد معلم متعامد، توضع على محوره الأفقي الفترات الزمنية، وعلى محوره العمودي قيم الظاهرة إنطلاقاً من معطيات إحصائية معينة، وعند إيصال نقاط الأزواج المرتبة (t, y) ، تظهر لنا طبيعة السلسلة الزمنية، التي قد تأخذ طبيعة تغيراتها عدة أشكال منها ما يلي :

1- تغيرات طويلة المدى : في مثل هذه الحالة يكون منحنى السلسلة غير متذبذب كثيراً، إذ تظهر عليه تموجات بسيطة متزايدة و متناقصة، على فترات زمنية طويلة ، وأهم ما يميزها أنها تستمر في اتجاه واحد لمدة طويلة من الزمن، سواء كان هذا الاتجاه متزايداً أو متناقصاً، وإذا حدث وأن تغير اتجاهها، فإنها ستبقى على هذا الاتجاه الجديد لفترة زمنية أخرى طويلة، ولعل أهم العوامل المؤثرة في منحنى مثل هذه التغيرات، خاصة إذا ما كانت الظاهرة تتعلق بكميات على المستوى الكلي، هو عدد السكان والتقدم التكنولوجي.

2- التغيرات الموسمية: في هذه الحالة تكون التغيرات دورية حسب أوقات معينة منتظمة، إذ يمكن أن تكون هذه التغيرات فصلية أي حسب الفصول الأربعة، أو شهرية . . . وقد تكون أسبوعية أو يومية، ومن العوامل الأكثر أهمية في إحداث التغيرات الموسمية، هي الطقس والعادات والتقاليد.

3- التغيرات الدورية: تتميز بوجود تغيرات دورية، ملاحظة على فترات زمنية طويلة نسبيا، تتكرر كل عدة سنوات، وهي تغيرات أطول من التغيرات الموسمية، وطول الدورة يكون غير معلوم وغير منتظم.

4- التغيرات العشوائية: وهي التغيرات غير المنتظمة التي تقع على الظاهرة، بسبب حالة طارئة غير متوقعة، وهي لا تحكمها قوانين أو قواعد معينة و بالتالي لا يمكن توقع حدوثها مسبقا وهي لا تستمر فترة زمنية طويلة، ومن أهم العوامل التي تؤدي إلى حدوث مثل هذه التغيرات، الإضرابات، الحرائق، الأحوال الجوية.

ثانيا : الخشونة : إن المنحنى البياني للسلسلة الزمنية، يمكن أن يأخذ عدة أشكال، إذ قد يكون شديد الإنكسار، وقد يكون ضعيف الإنكسار، أي أملس، والواقع أنه كلما كان المنحنى أملسا أي أكثر تمهدا، كلما أمكن أخذ فكرة واضحة، عن الاتجاه العام للسلسلة، و تقاس درجة إنكسار المنحنى عن طريق ما يسمى بمعامل الخشونة، الذي يعطى بالمعادلة التالية :

$$c = \frac{\sum_{t=1}^n (y_t - y_{t-1})^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad 2-8$$

حيث : y_t : قيمة الظاهرة في الفترة : t .

y_{t-1} : قيمة الظاهرة في الفترة : $t-1$

\bar{y} : الوسط الحسابي لقيم الظاهرة.

إذا كان : c كبيرا دل ذلك على شدة أنكسار منحنى السلسلة، وبالتالي صعوبة تحليل البيانات، وإذا ما كان صغيرا،

دل ذلك على أن السلسلة ملساء، أي منحناها ذو إنكسارات ضعيفة.

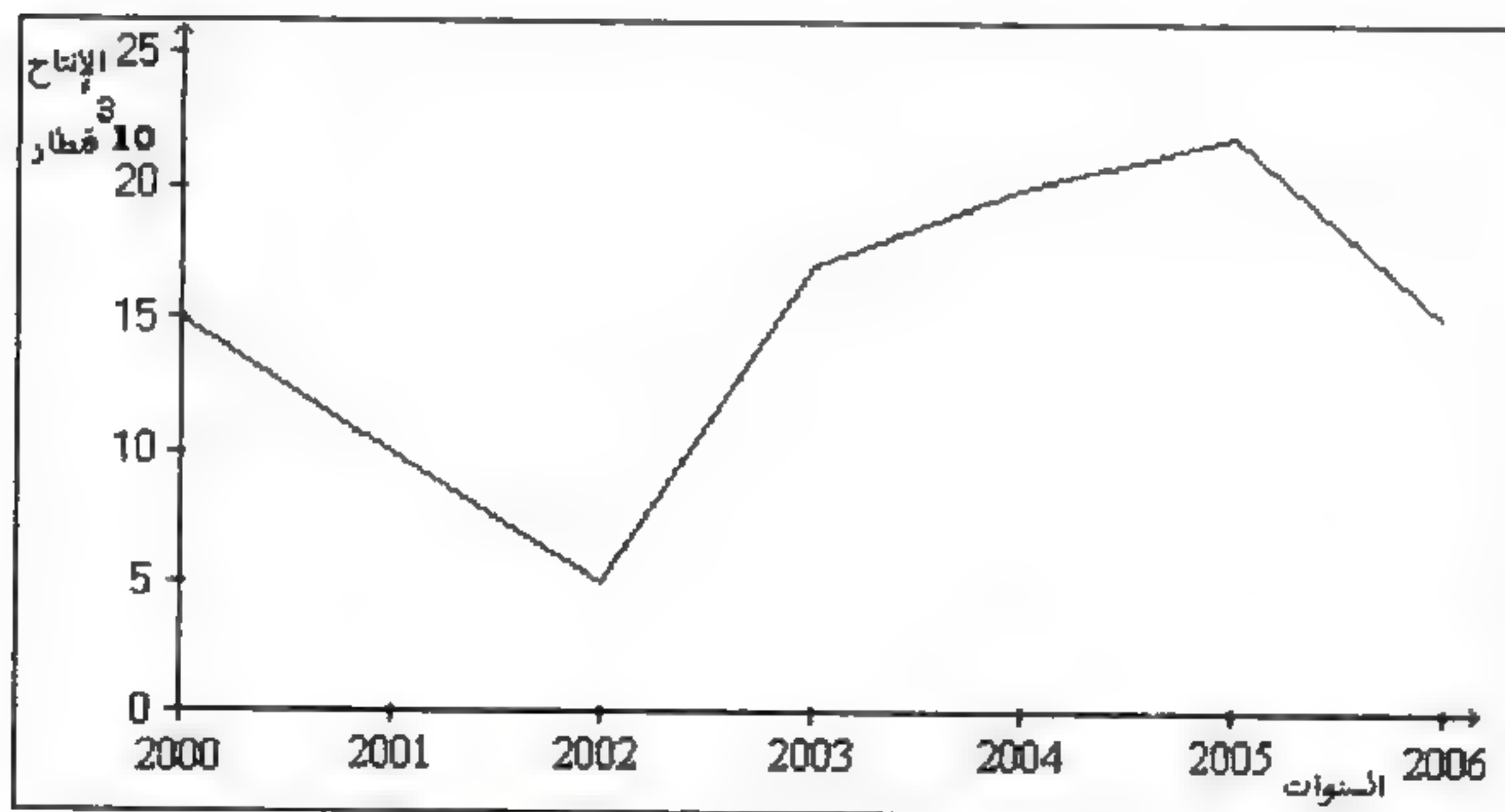
مثال 8-1: البيانات التالية تظهر تطور إنتاج الحبوب في مزرعة ما، خلال الفترة 2000-2006 بآلاف القناطير.

السنة	الإنتاج
2000	15
2001	10
2002	5
2003	17
2004	20
2005	22
2006	16

جدول 8-1

المطلوب: على معلم متعامد قدم هذه البيانات، ثم أوجد معامل خشونتها.
الجواب: المنحنى البياني التالي يظهر تطور إنتاج الحبوب، خلال الفترة: 2000-2006

تطور إنتاج الحبوب خلال الفترة 2000\2006



شكل 8-1

لإيجاد معامل الخشونة يتم استخدام المعادلة رقم 8-2 أعلاه، وبمساعدة الجدول

التالي :

economicrg
groups/economicrg

206
الباحث الاقتصادي

بوابة

Economicrg
researcher ate

السنة	t	y _t	(y _t - y _{t-1})	(y _t - y _{t-1}) ²	(y _t - \bar{y}) ²
2000	1	15	-	-	0
2001	2	10	-5	25	25
2002	3	5	-5	25	100
2003	4	17	12	144	4
2004	5	20	3	9	25
2005	6	22	2	4	49
2006	7	16	-6	36	1
مج		105		243	204

جدول 8-2

نجد:

$$\bar{y} = \frac{105}{7} = 15 \quad (10^3 \text{ قنطار})$$

$$c = \frac{243}{204} = 1.19$$

الوسط الحسابي للإنتاج :
معامل الخشونة:

تدل هذه القيمة أن السلسلة ليست خشنة بشكل كبير .
واضح أنه إذا ما كان معامل الخشونة كبيراً، فإنه يصعب تحديد
الاتجاه العام للسلسلة الزمنية، لذلك يلجأ إلى استخدام إحدى
الطرق الآتي ذكرها لتقدير اتجاهها.

ثالثاً: تقدير الاتجاه العام للسلسلة: يأخذ الاتجاه العام
للسلسلة عدة أشكال، منها الشكل الخطي وشبه الخطي،
والشكل الأسّي وغيره.

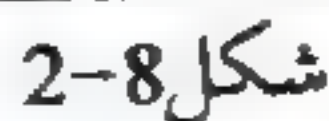
1- الشكل الخطي وشبه الخطي: يتم تحديد الاتجاه العام
في هذه الحالة بعدة طرق منها ما يلي:

أ- طريقة النقاط الوسطية : يتم على معلم متعامد تحسين
سلسلة البيانات، ثم نحدد النقاط p_i التي تمثل الإنكسارات
العلوية و نصل بينها، فنحصل على ما يسمى بخط السقف،
نحدد النقاط b_i التي تمثل الإنكسارات السفلية ونصل بينها،
فنحصل على ما يسمى بخط الأرضية، و إنطلاقاً من نقاط

مثال 8-2: أوجد خط الاتجاه العام بإستعمال طريقة النقاط الوسطية للبيانات التالية و التي تمثل تطور إنتاج إحدى المؤسسات بآلاف القناطر.

جدول 3-8

الإجابة : نرسم البيانات على معلم متعامد، و نقوم بالخطوات المحددة أعلاه فنحصل على الخط $A B$ الذي يمثل الاتجاه العام للسلسلة، كما هو واضح في الشكل 8-2.



فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

ب- طريقة المعدلات المتحركة:

تعريفه 8-2: إذا كانت لدينا البيانات :

$y_1, y_2, y_3 \dots y_n$ مأخوذة خلال الفترات :

$t=1, 2, 3, \dots, n$ ، على التوالي، فإن المعدلات المتحركة

بطول m ، لهذه السلسلة تعطى كما يلي :

المعدل المتحرك الأول :

$$y_1^* = \frac{y_1 + y_2 + y_3 + \dots + y_m}{m}$$

المعدل المتحرك الثاني :

$$y_2^* = \frac{y_2 + y_3 + y_4 + \dots + y_{m+1}}{m}$$

المعدل المتحرك الثالث :

$$y_3^* = \frac{y_3 + y_4 + y_5 + \dots + y_{m+2}}{m}$$

ويستمر إيجاد المعدلات المتحركة حتى نصل الى المعدل المتحرك الأخير، حيث يكتب :

$$y_{n-m+1}^* = \frac{y_{n-m+1} + y_{n-m+2} + y_{n-m+3} + \dots + y_n}{m}$$

ويكون عدد المعدلات المتحركة المحصل عليها في النهاية مساويا الى:

$$n-m+1$$

$$3-8$$

حيث n : عدد القيم . m طول المعدل المتحرك، ويفضل أن يكون عددا فرديا. و تكون السلسلة المحصل عليها بإستخدام المعدلات المتحركة أكثر ملوسة من السلسلة الأصلية، نستطيع من خلالها تخمين الاتجاه العام للسلسلة .

والجدير بالذكر هو أن طول الفترة التي يتعين إتخاذها أساسا لحساب المعدلات المتحركة، تتوقف على طبيعة البيانات ذاتها، ونشير الى أنه كلما كان طول المعدل المتحرك قصيرا كلما أمكن الحصول على خط

الجواب : باستخدام القاعدة المشار اليها أعلاه نحصل على قيم المعدلات المتحركة بطول 3 كما هي واضحة في العمود الثالث من الجدول التالي:

السنة	y_t	y^*
1998	6	
1999	4	6.66
2000	10	7.33
2001	8	9.66
2002	11	8.33
2003	6	10
2004	13	10
2005	11	13
2006	15	

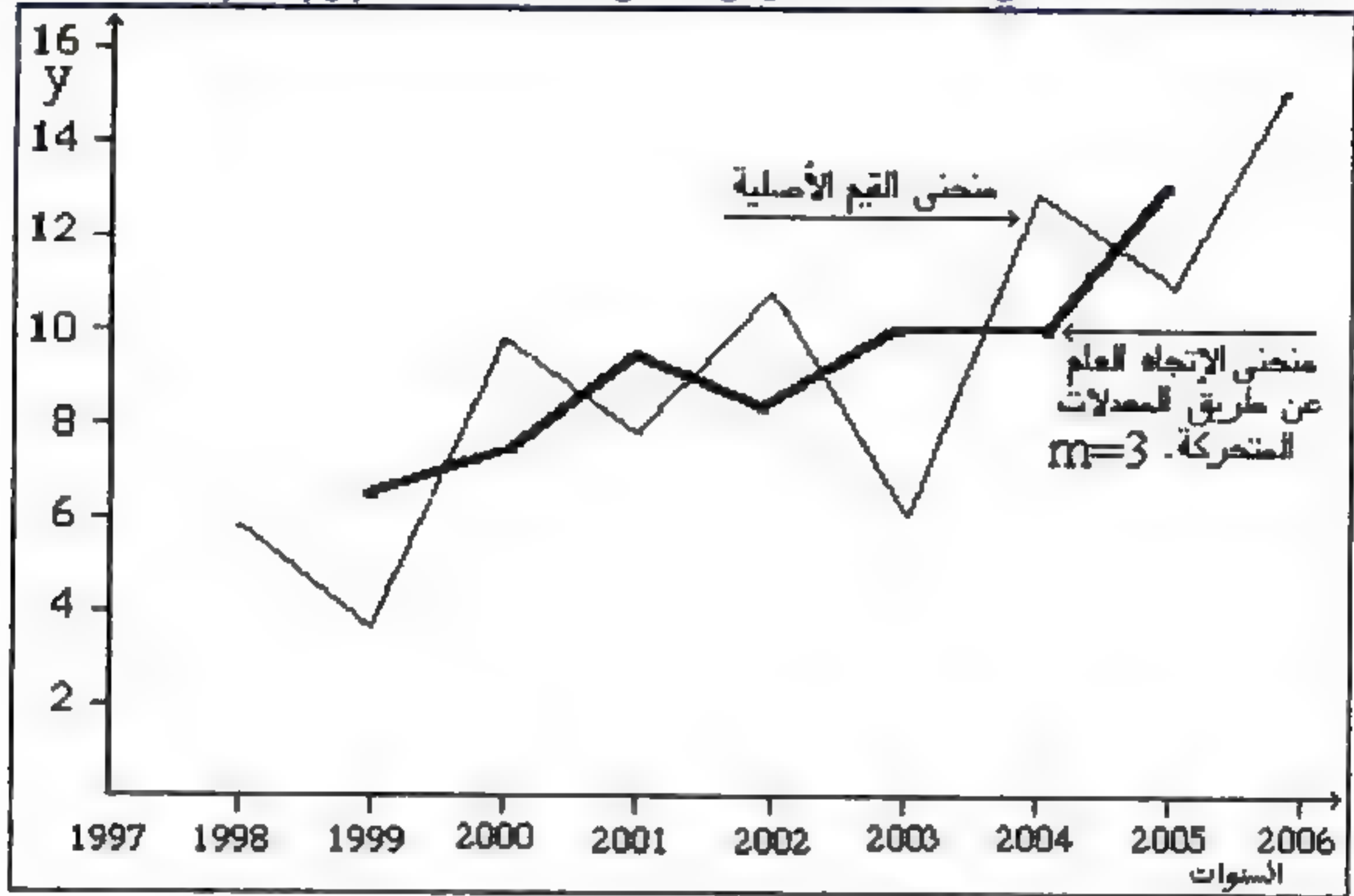
جدول 4-8

$$y^*_1 = (6+4+10)/3 = 6.66$$

$$y^*_2 = (4+10+8)/3 = 7.33$$

و بالمثل تم إيجاد بقية المعدلات المتحركة.

عند رسم البيانات الأصلية والبيانات المحصل عليها باستخدام طريقة المعدلات المتحركة، يظهر منحنى الاتجاه العام للسلسلة بوضوح أكثر، إذ أن إنكسارات السلسلة الأصلية تختفي وتظهر سلسلة المعدلات المتحركة أقل خشونة، وذلك ما يظهره الشكل الموالي:



شكل 3-8

ج- **طريقة المعدلات الجزئية** : في هذه الطريقة يتم تقسيم السلسلة الى مجموعة من السلاسل الجزئية، بطول m ، ثم يتم إيجاد متوسطات هذه السلاسل . تجسد هذه المتوسطات على معلم متعامد ويوصل بينها فنحصل بذلك على منحنى الاتجاه العام، وهو منحنى أقل خشونة من المنحنى الأصلي .

تعريفه 3-8: إذا كانت لدينا السلسلة : $y_1, y_2, y_3, \dots, y_n$ ، فإن معدلاتها المتدرجة بطول m تعرف كما يلي :

المعدل الأول:

$$\bar{y}_1 = \frac{y_1 + y_2 + y_3 + \dots + y_m}{m}$$

المعدل الثاني:

$$\bar{y}_2 = \frac{y_{m+1} + y_{m+2} + y_{m+3} + \dots + y_{m+m}}{m}$$

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018
المعدل الثالث:

$$\bar{y}_3 = \frac{y_{2m+1} + y_{2m+2} + y_{2m+3} + \dots + y_{2m+m}}{m}$$

وبالمثل تحسب بقية المعدلات.

حيث m : طول السلسلة الجزئية، ويفضل أن يكون فرديا ومن قواسم N .

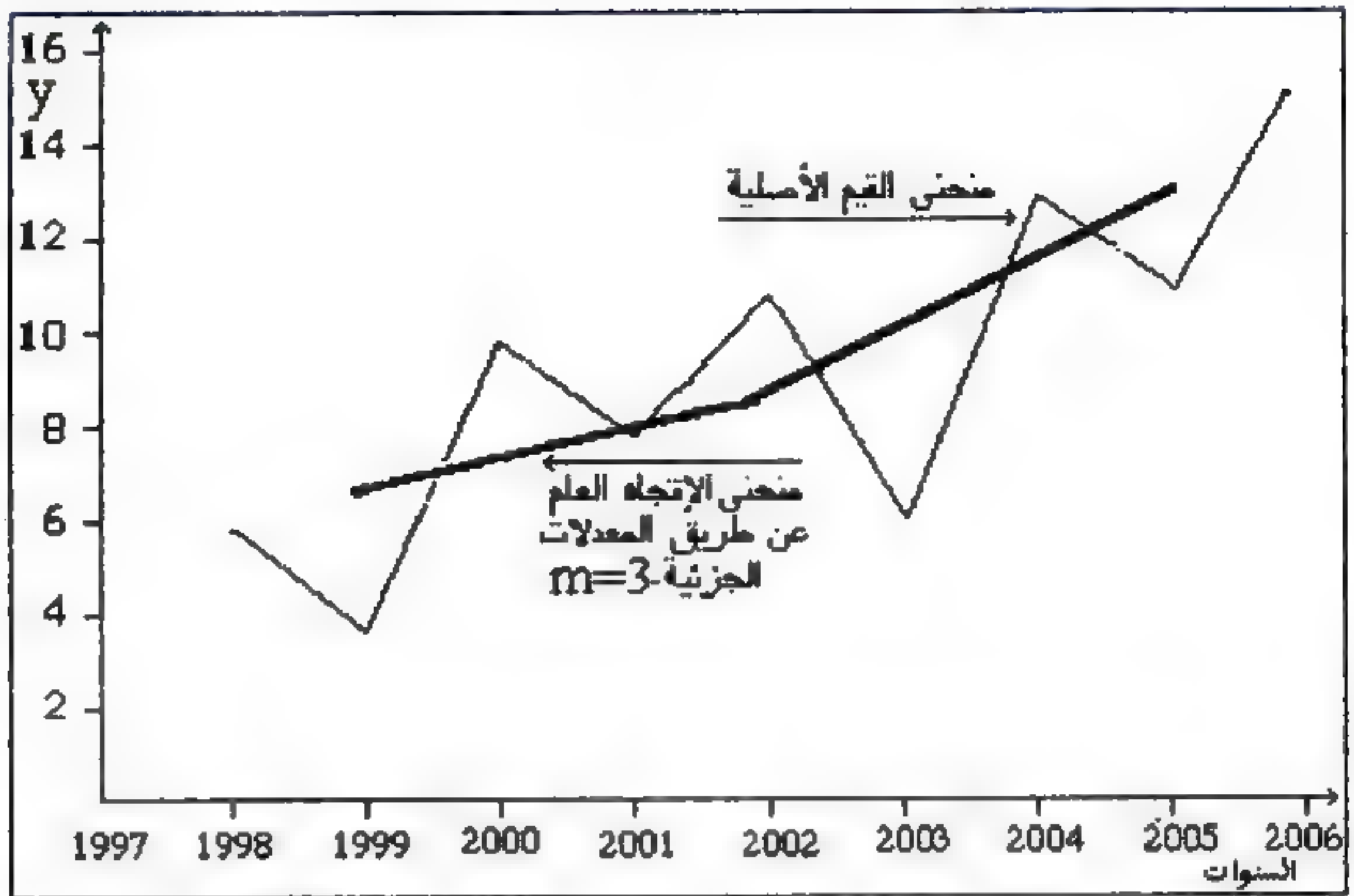
مثال 4-8: باستخدام طريقة المعدلات الجزئية بطول 3، حدد خط الاتجاه العام لسلسلة المثال 2-8.

الإجابة: باستخدام القاعدة أعلاه نحصل على جدول المعدلات الجزئية كما هي واردة في الجدول التالي:

السنة	y_t	\bar{y}_i
1998	6	
1999	4	6.66
2000	10	
2001	8	
2002	11	8.33
2003	6	
2004	13	
2005	11	13
2006	15	

جدول 5-8

عند تجسيم البيانات المحصل عليها نجد الاتجاه العام للسلسلة، كما يحدده الشكل 4-8



شكل 4-8

يلاحظ أن المنحنى المحصل عليه، يكاد يكون مستقيماً، وهو يحدد الاتجاه العام للسلسلة بكل وضوح.

د- طريقة المتوسطات النصفية: في هذه الطريقة يتم فصل البيانات الى قسمين متساويين، نوجد الوسط الحسابي لبيانات القسم الأول، ونوجد كذلك الوسط الحسابي لبيانات القسم الثاني، نحدد نقطتي الوسطين على معلم متعامد ونصل بينهما فنحصل على خط مستقيم، يحدد الاتجاه العام للسلسلة. تستخدم هذه الطريقة فيما لو استنتج الباحث أنه يمكن تمثيل البيانات عن طريق خط مستقيم.

هـ- طريقة المربعات الصغرى: في هذه الحالة يفترض أن تكون السلسلة دالة خطية في الزمن، من الشكل :

$$y_t = a + bt \quad 4-8$$

حيث : $t = 0, 1, 2, \dots, n-1$

بحيث نعطي الفترة الأولى : $t=0$ ، الفترة الثانية : $t=1$... الفترة الأخيرة : $t=n-1$

وتقدر معالم الدالة: b و a بطريقة المربعات الصغرى، عن طريق معادلتين مشابھتين لمعادلات تقدير معالم خط الانحدار رقم 6-7 و 7-7، أي:

خطأ! الإشارة المرجعية غير معروفة. 5-8

$$a = \bar{y} - b \bar{t}$$

6-8

حيث:

$$\bar{y} = \frac{\sum_{t=0}^{n-1} y_t}{N}$$

7-8

$$\bar{t} = \frac{\sum_{t=0}^{n-1} t}{N}$$

8-8

مثال 5-8: البيانات التالية تعرض تطور صادرات المحروقات خلال الفترة 1998-2006 بملايين الدينارات لدولة ما.

المطلوب: أوجد خط الاتجاه العام لهذه البيانات، وأوجد القيم المقدرة (القيم الاتجاهية) للصادرات:

الصادرات	السنة
35	1998
39	1999
41	2000
46	2001
47	2002
50	2003
52	2004
56	2005
60	2006

جدول 6-8

بأخذ سنة 1998 كنقطة أصلية، وبتطبيق مبدأ المربعات الصغرى في إيجاد معالم الدالة: $y_t = a + b_t$. ومن خلال الجدول التالي نجد:

السنة	y_t	t	$(t - \bar{t})(y_t - \bar{y})$	$(t - \bar{t})^2$	القيم الاتجاهية
1998	35	0	49.33	16	35.53
1999	39	1	25.00	9	38.48
2000	41	2	12.67	4	41.43
2001	46	3	01.33	1	44.38
2002	47	4	00.00	0	47.33
2003	50	5	02.67	1	50.28
2004	52	6	09.33	4	53.23
2005	56	7	26.00	9	56.18
2006	60	8	50.67	16	59.13
مج	426	36	177.00	60	

جدول 7-8

$$\sum_{t=0}^8 t = 36 \Rightarrow \bar{t} = 4$$

$$\sum_{t=0}^8 y_t = 426 \Rightarrow \bar{y} = 47.33$$

$$\sum_{i=0}^8 (y_t - \bar{y})(t - \bar{t}) = 177$$

$$\sum_{i=0}^8 (t - \bar{t})^2 = 60$$

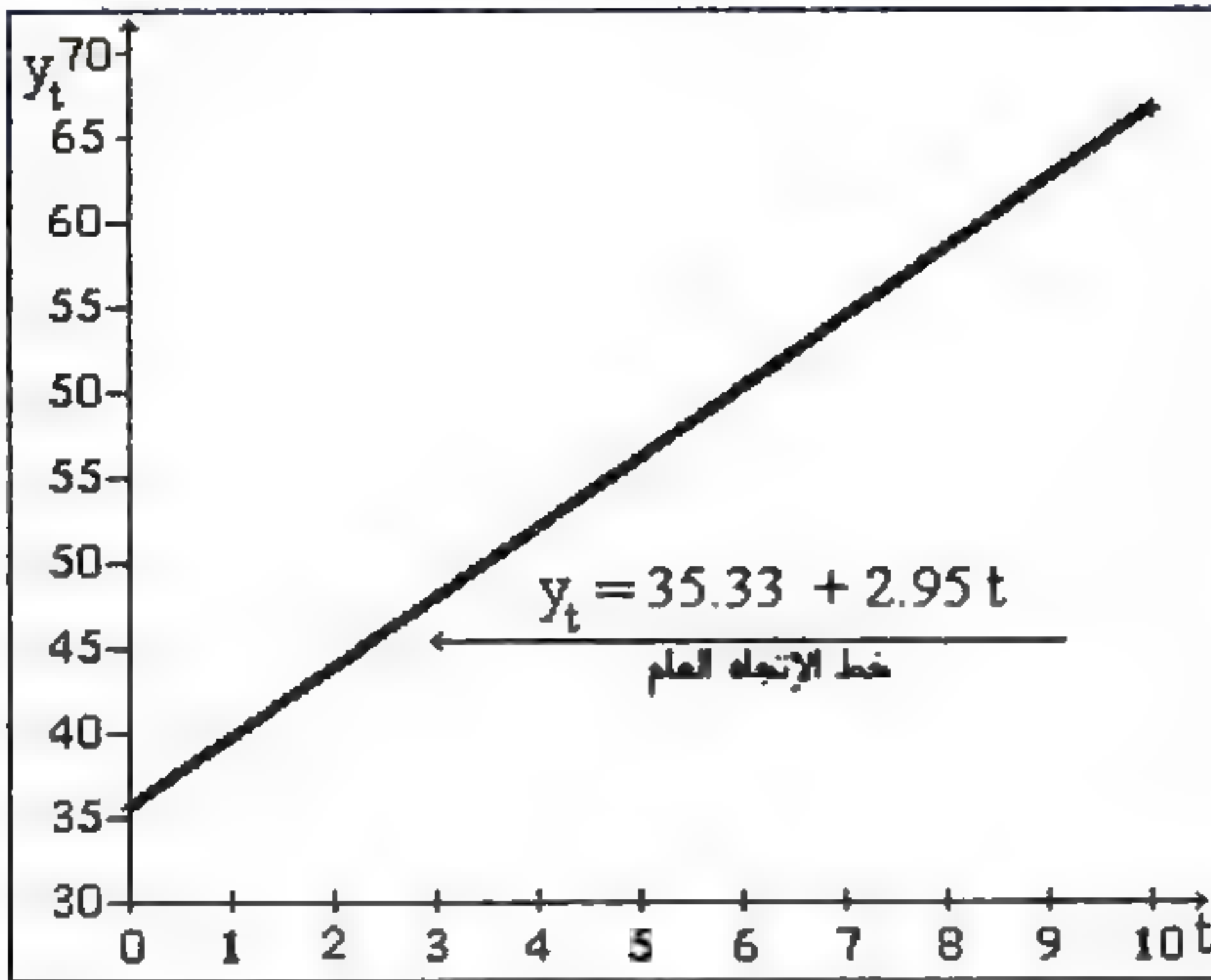
$$a=35.53 \quad , \quad b=2.95$$

ومنه يكون :

وبالتالي تكون معادلة خط الاتجاه العام المقدرة بطريقة المربعات الصغرى كما يلي:

$$y_t = (35.53 + 2.95t) \quad \text{مليون دينار} \quad 9-8$$

ويظهر خط الاتجاه العام المقدر في الشكل 5-8.



شكل 8-5 .

يمكن إيجاد نفس القيم الاتجاهية الموجودة في العمود الأخير من الجدول 8-7. بما يسمى بالطريقة المختصرة، وذلك يجعل السنة الوسيطة هي نقطة الأصل في حالة ما إذا كان N فرديا، وتأخذ السنوات التي تسبقها ترتيبا سالبا و السنوات التي تليها ترتيبا موجبا، وتشكل هذه الترتيب قيم المتغير المستقل، ففي مثالنا السابق، السنة الوسيطة هي سنة 2002 لذلك فهي تشكل نقطة الأصل و فيها يكون $x_0 = 0$ ويكون في السنة التي تليها $x_1 = 1$ وهكذا بالنسبة للبقية، أما السنوات التي تسبقها فتأخذ إشارات سالبة، و يوضح ذلك الجدول 8-8.

السنة	y_t	x_t
1998	35	-4
1999	39	-3
2000	41	-2
2001	46	-1
2002	47	0
2003	50	1
2004	52	2
2005	56	3
2006	60	4
مج	426	0

جدول 8-8

من خلال المعادلتين الطبيعيتين 10-8 و 11-8:

$$\sum_{t=1}^n y_t = b \sum_{t=1}^n x_t + Na \quad 10-8$$

$$\sum_{t=1}^n y_t x_t = b \sum_{t=1}^n x_t^2 + a \sum_{t=1}^n x_t \quad 11-8$$

$$\sum_{t=1}^n y_t = Na \quad 12-8 \quad \text{نجد :}$$

$$\sum_{t=1}^n y_t x_t = b \sum_{t=1}^n x_t^2 \quad 13-8$$

$$\sum_{t=1}^n x_t = 0 \quad \text{ذلك لأن :}$$

$$a = \frac{\sum_{t=1}^n y_t}{N} \quad 14-8 \quad \text{و منه يكون :}$$

$$b = \frac{\sum_{t=1}^n y_t x_t}{\sum_{t=1}^n x_t^2} \quad 15-8 \quad \text{و}$$

بتطبيق ذلك على المثال السابق و بالاستعانة بالجدول التالي :

السنة	y_t	x_t	$y_t x_t$	x_t^2	القيمة الاتجاهية
1998	35	-4	-140	16	35.53
1999	39	-3	-117	9	38.48
2000	41	-2	-82	4	41.43
2001	46	-1	-46	1	44.38
2002	47	0	0	0	47.33
2003	50	1	50	1	50.28
2004	52	2	104	4	53.23
2005	56	3	168	9	56.18
2006	60	4	240	16	59.13
مج	426		177	60	

جدول 8-9

$$b = 177/60 = 2.95 \quad a = 426/9 = 47.33 \quad \text{نجد :}$$

و بالتالي تكون الدالة الزمنية على النحو :

$$y_t = 47.33 + 2.95 x_t$$

في العمود الأخير من الجدول 8-9 تظهر لنا القيم الاتجاهية و هي نفس القيم الاتجاهية الواردة في العمود الأخير من الجدول 8-7، ويظهر أن للمعادلتين نفس الميل وهو : $b=2.95$.

هذا في حالة ما إذا كان N فرديا بحيث تكون السنة الوسيطة سنة وحيدة، أما إذا كان N زوجيا فإنه يتم أخذ المتغير المستقل على أساس نصف سنوي، لإيجاد القيم الاتجاهية، وذلك كما في المثال التالي :

مثال 8-6 : البيانات التالية تظهر تطور إنتاج الحبوب في مزرعة ما خلال الفترة 1999-2006 بآلاف القناطير.

السنة	1999	2000	2001	2002	2003	2004	2005	2006
الإنتاج	5	7	7	10	13	15	18	20

جدول 8-10

المطلوب : أوجد القيم الاتجاهية للإنتاج بالطريقة المختصرة.

بما أن N زوجي فإنه لا توجد سنة وسيطة وحيدة، لذلك نأخذ قيم المتغير المستقل على أساس وحدات نصف سنوية، وتكون قيم x كما في الجدول التالي:

السنة	y_t	x_t
1999	5	-7
2000	7	-5
2001	7	-3
2002	10	-1
2003	13	1
2004	15	3
2005	18	5
2006	20	7
مج		0

جدول 8-11

باستخدام الطريقة المختصرة نوجد فقط : $y_t x_t$ و x_t^2 وذلك

من خلال الجدول 8-12:

السنة	y_t	x_t	$y_t x_t$	x_t^2	القيم الاتجاهية
1999	5	-7	-35	49	4.08
2000	7	-5	-35	25	6.31
2001	7	-3	-21	9	8.54
2002	10	-1	-10	1	10.76
2003	13	1	13	1	12.99
2004	15	3	45	9	15.21
2005	18	5	90	25	17.44
2006	20	7	140	49	19.67
مج		0	187	168	

جدول 8-12

باستخدام المعادلتين 8-14 و 8-15 نجد :

$$a = 11.88$$

$$b = 1.1$$

ومنه تكون المعادلة الزمنية بالطريقة المختصرة هي :

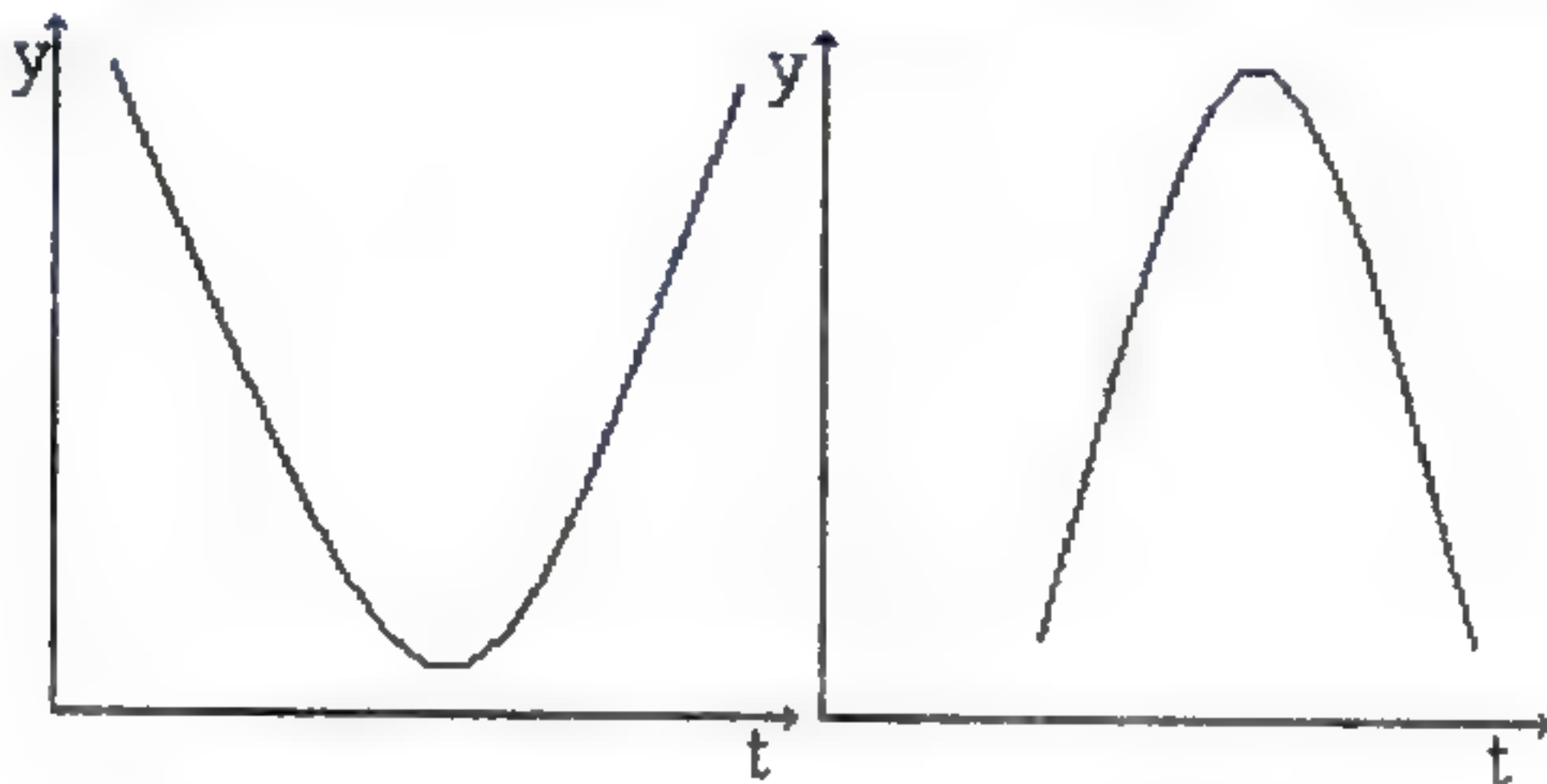
$$y_t = 11.88 + 1.1 x_t$$

و تظهر القيم الاتجاهية في العمود السادس من الجدول أعلاه، وهي نفس القيم الاتجاهية التي نحصل عليها فيما لو استخدمنا الطريقة العادية، إذ تكون معادلة الاتجاه العام بهذه الطريقة (الطريقة العادية) على النحو :

$$y_t = 4.08 + 2.23 t$$

2- الشكل غير الخطي للاتجاه العام : يمكن أن يأخذ الاتجاه العام للسلسلة الزمنية شكلا غير خطي، كشكل دالة القطع المكافئ أو شكل الدالة الأسية وغيرها، وسنتناول في هذا البند كيفية تقدير الاتجاه العام للظواهر فيما لو ظهر شكل انتشار قيمها على الزمن بشكل يوحي بأن دالتها الزمنية تكون في صيغة دالة قطع مكافئ أو صيغة دالة أسية.

1- صيغة دالة القطع المكافئ : يأخذ منحنى دالة القطع المكافئ أحد الشكلين التاليين:



شكل 7-8

شكل 8-6

ونظرا لإستحالة إيجاد الدالة الحقيقية، فإنه يتم تقدير الاتجاه العام لها عن طريق المعادلة التالية:

$$y_t = a + bt + ct^2$$

16-8

ويتم تقدير معالم هذه الدالة باستخدام طريقة المربعات الصغرى، وذلك بتحويلها الى معادلة خطية، بإجراء التحويلات التالية:

$$t=x, \quad t^2=z$$

17-8

و بناء على ذلك تصبح المعادلة الجديدة على النحو :

$$y_t = a + bx_t + cz_t$$

18-8

وكما هو واضح، فلإن المعادلة المحصل عليها هي معادلة إنحدار ثلاثي، لذلك يتم تقدير معالمها: a ، b ، c عن طريق المعادلات التالية: 19-8، 20-8، 21-8:

$$b = \frac{\sum_{i=1}^n y_i x_i \sum_{i=1}^n z_i^2 - \sum_{i=1}^n y_i z_i \sum_{i=1}^n x_i z_i}{\sum_{i=1}^n x_i^2 \sum_{i=1}^n z_i^2 - (\sum_{i=1}^n x_i z_i)^2}$$

19-8

$$c = \frac{\sum_{i=1}^n y_i z_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i x_i \sum_{i=1}^n x_i z_i}{\sum_{i=1}^n x_i^2 \sum_{i=1}^n z_i^2 - (\sum_{i=1}^n x_i z_i)^2} \quad 20-8$$

$$a = \bar{y} - b\bar{x} - \bar{z}c \quad 21-8$$

و هي نفس المعادلات : 6-34، 6-35، 6-36 المستخدمة في الإنحدار الثلاثي في الفصل السابق.

بعد الحصول على المعالم المقدرة : a ، b ، c ، يتم تحويل المعادلة المحصل عليها الى أصلها كما هي واردة في المعادلة رقم 8-16.

مثال: 8-7 : البيانات التالية تظهر تطور السواردات الغذائية خلال الفترة 1998-2006، بملايير الدينارات.

السنة	القيمة
1998	50
1999	100
2000	200
2001	350
2002	400
2003	350
2004	200
2005	100
2006	50

جدول 8-13

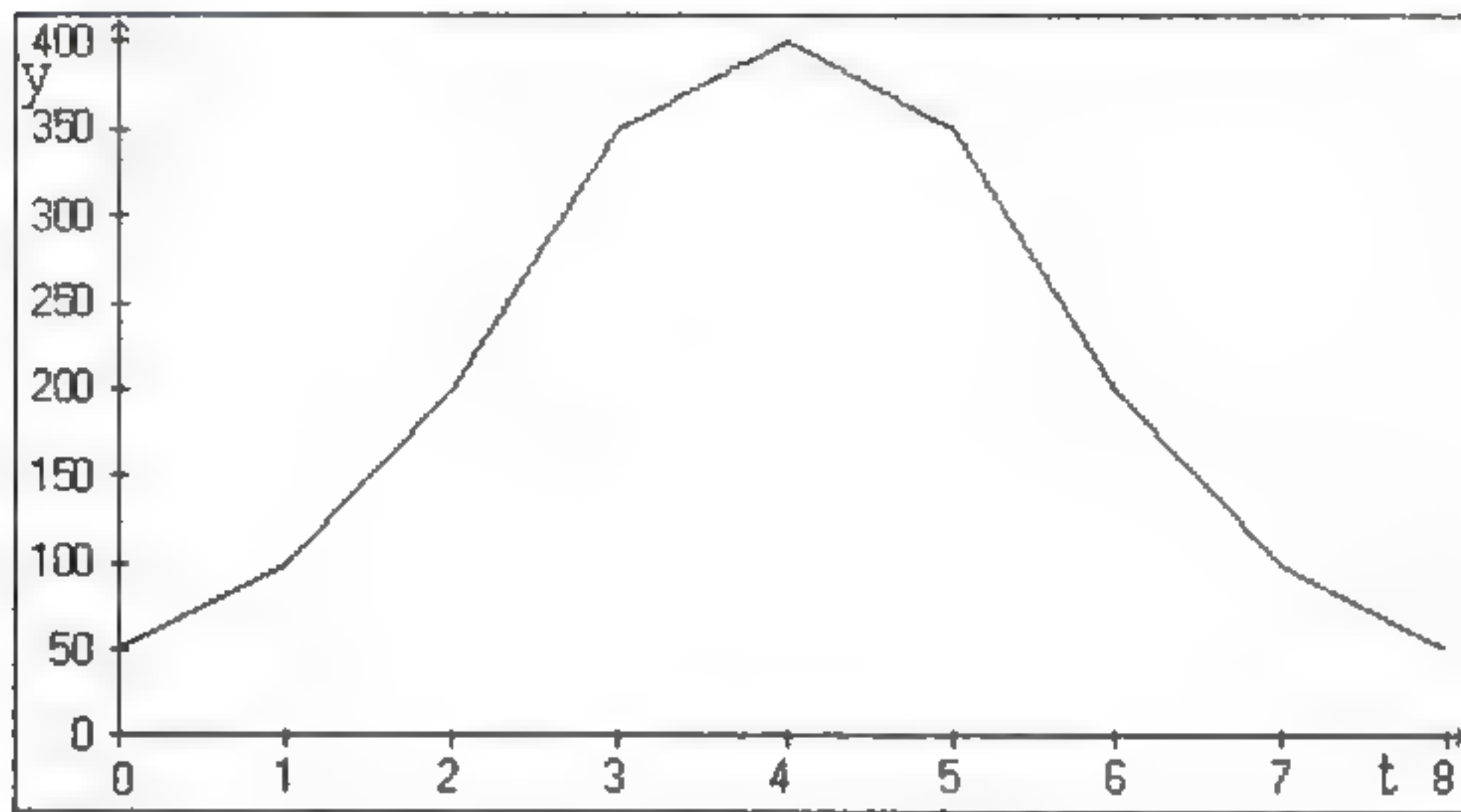
المطلوب : 1- قدم هذه البيانات على معلم متعامد. ماذا تستنتج؟

2- إعط المعادلة العامة للشكل المحصل عليه، ثم قدرها.

الإجابة : تقدم البيانات على معلم متعامد:

تطور الواردات الغذائية خلال الفترة 1998-2006، بملايير الدينارات

t=0 سنة 1998



شكل 8-8

يلاحظ أن الشكل المحصل عليه هو شكل دالة قطع مكافئ، وبالتالي فإن شكل دالته سوف يكون على نحو الدالة رقم 8-16، أي:

$$y_t = a + bt + ct^2$$

بإجراء التحويلات اللازمة، لتصبح الدالة خطية يمكن تقديرها بطريقة المربعات الصغرى، نحصل على معادلة من الشكل 8-18 أي :

$$y_t = a + bx_t + cz_t$$

$$x_t = t, z_t = t^2$$

حيث :

ونجد معالم الدالة بتطبيق المعادلات 8-19، 8-20، 8-21 بمساعدة الجدول التالي:

t=x	t ² =z	y _t	y _t x _t	y _t z _t	z _t x _t	x _t ²	z _t ²	القيم الاتجاهية
0	0	50	0	0	0	0	0	1.5
1	1	100	100	100	1	1	1	148.5
2	4	200	400	800	8	4	16	253.8
3	9	350	1050	3150	27	9	81	317.4
4	16	400	1600	6400	64	16	256	339.1
5	25	350	1750	8750	125	25	625	319.2
6	36	200	1200	7200	216	36	1296	257.5
7	49	100	700	4900	343	49	2401	154.0
8	64	50	400	3200	512	64	4096	8.9
36	204	1800	7200	34500	1296	204	8772	مج

جدول 8-14

$$b=167.89$$

نجد:

$$c=-20.87$$

$$a=1.53$$

ومنه تكون معادلة الانحدار المقدرة على أساس المعادلة رقم 8-18 هي :

$$y_t = 1.53 + 167.89x_t - 20.87z_t$$

بتحويلها الى أصلها كما في المعادلة 8-16 نحصل على دالة القطع المكافيء المقدرة كما هي أدناه :

$$y_t = 1.53 + 167.89t - 20.87t^2$$

تكون القيم الاتجاهية تقريبية كما هي واضحة في العمود الأخير من الجدول 8-14.

كما سبقت الإشارة تستخدم هذه الطريقة في إيجاد الاتجاه العام للبيانات التي يعطى شكل انتشارها شكلاً مشابهاً للقطع المكافيء، وعملياً نصادف الكثير من البيانات من هذا النوع، كتطور واردات دولة ما من الجيوب، فقد نجد أن وارداتها تزايدت مع البداية، ولكن باتباع سياسة الاعتماد على النفس وإحلال المنتج المحلي محل الواردات، فإن الواردات تبدأ في التناقص بعد أن تصل مستوى أعظمي معين، وذلك بوتائر سلبية، الشيء الذي يجعل اتجاهها العام يشبه القطع المكافيء.

2- صيغة الدالة الأسية : في بعض الأحيان يتخذ الاتجاه العام للسلسلة الزمنية شكل الدالة الأسية، وذلك خاصة بالنسبة للمتغيرات القابلة للنمو بوتائر ثابتة على امتداد فترات زمنية معينة، و تكون الصيغة العامة للدالة الأسية على النحو التالي :

$$y_t = a.b^t . \varepsilon_t$$

8-22

حيث : a ، b ثوابت، ε_t : حد الخطأ وهو متغير عشوائي. ويتم تقديرها عن طريق المعادلة التالية :

$$y_t = a.b^t$$

8-23

حيث : a : هي قيمة تقاطع المنحنى بالإحداثي العمودي، b : معامل نمو الدالة.

و باستخدام طريقة المربعات الصغرى يتم تقدير معالم هذه الدالة وذلك بعد تحويلها الى دالة خطية بإدخال اللوغاريتم على طرفيها، أي :

$$\text{Log } b \cdot \text{Log } y_t = \text{Log } a \cdot b^t = \text{Log } a + t \quad 24-8$$

وبإجراء التحويلات التالية :

$$\text{Log } y_t = y_t^*$$

$$\text{Log } a = a^*$$

$$\text{Log } b = b^*$$

نحصل على الدالة التالية:

$$y_t^* = a^* + b^* t \quad 25-8$$

باستخدام طريقة المربعات الصغرى نحصل على المعادلتين التقديريتين التاليتين:

$$b^* = \frac{\sum_{t=0}^{n-1} (y_t^* - \bar{y}^*)(t - \bar{t})}{\sum_{t=0}^{n-1} (t - \bar{t})} \quad 26-8$$

$$a^* = \bar{y}^* - b^* \bar{t} \quad 27-8$$

بعد إيجاد معالم الدالة 25-8 يتم تحويلها الى الأصل حسب المعادلة 23-8. القيمة b كما هي في المعادلة 23-8 عبارة عن نسبة القيمة التقديرية في فترة معينة الى قيمتها في الفترة السابقة لها، أي :

$$b = \frac{y_t}{y_{t-1}} \quad 28-8$$

$$y_t \text{ تكتب على النحو: } y_t = y_{t-1} + (y_t - y_{t-1}) \text{ أو :}$$

$$y_t = y_{t-1} + \Delta y_{t-1} \quad 29-8$$

بتعويض 29-8 في المعادلة 28-8 نجد:

$$b = \frac{y_{t-1} + \Delta y_{t-1}}{y_{t-1}} = 1 + \frac{\Delta y_{t-1}}{y_{t-1}} \quad 30-8$$

معلوم أن الطرف الأخير من المعادلة 8-30 هو القيمة التقديرية لوتيرة النمو (معدل النمو). أنظر البند الثاني من الفصل 9)، ونرمز له بالحرف k ، لذلك يمكن أن نكتب b على النحو التالي:

$$b=1+k$$

31-8

و النتيجة هي أن نمو قيمة المتغير التابع بالنسبة للزمن، بموجب الدالة الأسية يكون في شكل متوالية هندسية، و تكون الزيادة النسبية ثابتة، والجدول التالي يظهر ذلك:

t	$y_t = a \cdot b^t$	Y_t / y_{t-1}
0	a	-
1	a.b	b
2	a.b ²	b
3	a.b ³	b
.	.	.
.	.	.
.	.	.
n-1	a.b ⁿ⁻¹	b

جدول 8-15

من الجدول يظهر إذن أن الرقم القياسي التقديري ثابت على طول الفترة المدروسة و يساوي القيمة b .

تعتبر الدالة الأسية جد هامة في التطبيقات الاقتصادية و الإجتماعية، خاصة بالنسبة للظواهر التي تنمو بمعدلات نمو ثابتة، كتطور السكان وتطور الإنتاج. مثال 8-8: البيانات التالية تظهر تطور إنتاج النفط في أحد الحقول بآلاف البراميل:

السنة	الإنتاج
2000	4
2001	8
2002	20
2003	40

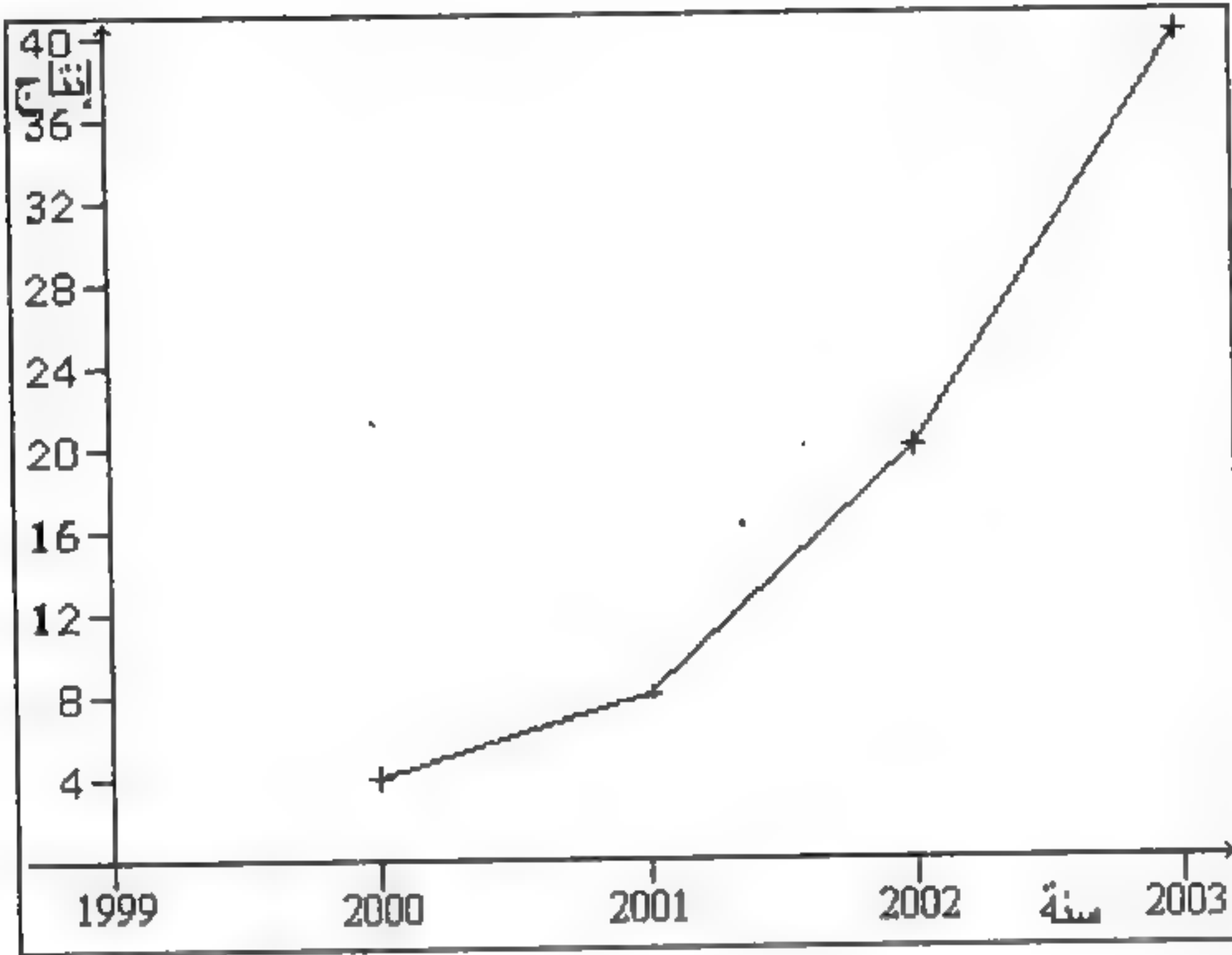
جدول 8-16

المطلوب: ارسم شكل الانتشار واستنتج الدالة الزمنية التي يخضع لها تطور الإنتاج، ثم قدرها.

الإجابة:

شكل الانتشار:

تطور إنتاج النفط في أحد الحقول.



شكل 8-9

من شكل الانتشار يظهر بأن الدالة الزمنية لتطور الإنتاج تأخذ شكل الدالة الأسية كما هي معرفة أعلاه أي :

$$y_t = a \cdot b^t$$

حيث y_t : قيمة الإنتاج في الفترة t

و يتم تقدير معالمها a و b بتحويلها الى شكل خطي لنتمكن من تقديرها باستخدام طريقة المربعات الصغرى و ذلك كما هو واضح أعلاه، وعليه يتعين علينا إجراء التحويلات التالية :

بجعل:

$$\text{Log } y_t = y_t^*$$

$$\text{Log } a = a^*$$

$$\text{Log } b = b^*$$

تكون الدالة الخطية المطلوب تقدير معالمها بطريقة المربعات الصغرى على الشكل:

$$y_t^* = a^* + b^* t$$

بتطبيق المعادلات 26-8 و 27-8 و من خلال الجدول التالي نجد :

السنة	t	y _t	y _t [*] = Log y _t	(y _t [*] - \bar{y}^*)(t - \bar{t})	(t - \bar{t}) ²
2000	0	4	0.60	0.75	2.25
2001	1	8	0.90	0.10	0.25
2002	2	20	1.30	0.10	0.25
2003	3	40	1.60	0.75	2.25
مج	6		4.40	1.70	5

جدول 8-17

$$\text{Log } b = b^* = 1.7/5 = 0.34$$

$$\text{Log } a = a^* = 0.59$$

و تكون :

$$b = 10^{0.34} = 2.19$$

$$a = 10^{0.59} = 3.89$$

و بالتالي تكون الدالة الزمنية المقدرة على النحو :

$$y_t = 3.89 \cdot 2.19^t$$

أما القيم الاتجاهية للدالة فهي:

$$40.89, 18.66, 8.52, 3.89$$

و هي قيم قريبة جدا من القيم الحقيقية.

رابعاً: تحديد المركبات الموسمية وحذف تغيراتها : كما

سبقت الإشارة فإن السلسلة الزمنية تحتوي على تغيرات الاتجاه العام وعلى التغيرات الموسمية والدورية وكذا التغيرات العشوائية، ويكون من المفيد، تحديد مركبات الموسم، وحذف التغيرات الموسمية، ذلك لأن هدف الباحث هو التعرف على القيمة الحقيقية للظاهرة مستبعداً منها تأثير التقلبات الموسمية، تجنباً لتضليل الإرتفاع أو الإنخفاض المؤقت في قيمة الظاهرة.

ولأجل ذلك هناك عدة طرق نتناول منها الطريقتين التاليتين :

1- طريقة النسبة إلى الاتجاه العام : كما هو واضح من

إسم هذه الطريقة، فإنه يتم الإعتماد على نسبة القيم الحقيقية الى القيم الاتجاهية المحصل عليها إنطلاقاً من إحدى طرق تحديد خط الاتجاه العام وذلك بإتباع الخطوات التالية:

أ- تحديد خط الاتجاه العام.

ب- إيجاد النسبة المئوية بقسمة كل قيمة حقيقية على القيمة الاتجاهية المقدرة المقابلة.

ج- إيجاد حاصل قسمة مجموع النسب المحصل عليها من الخطوة ب على عدد السنوات، فنحصل بذلك على معدلات تسمى بالمعاملات الموسمية أو المركبات الموسمية.

د- نقسم كل قيمة حقيقية على المعامل الموسمي المقابل لها و نضربه في 100 فنحصل على قيم الظاهرة محذوفة التغيرات الموسمية.

مثال 8- 9: أوجد القيم المحذوفة التغيرات الموسمية للبيانات التالية بطريقة النسبة الى الاتجاه العام.

شهر/سنة	1	2	3	4	5	6	7	8	9	10	11	12
1992	4	5	6	7	6	7	8	3	6	6	5	6
1993	6	7	8	6	9	10	9	4	8	11	12	14
1994	10	12	13	15	15	15	16	9	14	15	16	17

جدول 8-18

الجواب:

نلاحظ أن القيم الموسمية في هذا المثال عبارة عن قيم شهرية.

أ- نوجد خط الاتجاه العام باستخدام طريقة المربعات الصغرى إنطلاقاً من القيم الشهرية حسب ترتيبها بمساعدة الجدول 8-19، حيث نجد الدالة الخطية للإتجاه العام كما يلي :

$$y_t = 3.22 + 0.34t$$

حيث t ترتيب الشهر ($t = 1, 2, 3, \dots, 36$).

و تظهر القيم الاتجاهية في العمود الخامس من الجدول 8-19.

t	y_t	$(t-t)(y_t - y)$	$(t-t)^2$	قيم الاتجاه	t	y_t	$(t-t)(y_t - y)$	$(t-t)^2$	قيم الاتجاه
1	4	59.28	306.25	3.56	19	9	-6.22	0.25	9.61
2	5	73.33	272.25	3.89	20	4	-8.17	2.25	9.95
3	6	53.39	240.25	4.23	21	8	-3.61	6.25	10.29
4	7	35.44	210.25	4.57	22	11	5.41	12.25	10.62
5	6	46.50	182.25	4.90	23	12	11.50	20.25	10.96
6	7	30.56	156.25	5.24	24	14	25.06	30.25	11.29
7	8	16.61	132.25	5.58	25	10	3.61	42.25	11.63
8	3	67.67	110.25	5.91	26	12	19.17	56.25	11.97
9	6	32.72	90.25	6.25	27	13	30.22	72.25	12.30
10	6	29.28	72.25	6.58	28	15	52.78	90.25	12.64
11	5	33.33	56.25	6.92	29	15	58.33	110.25	12.98
12	6	22.39	42.25	7.26	30	15	63.89	132.25	13.31
13	6	18.34	30.25	7.59	31	16	81.94	156.25	13.65
14	7	11.00	20.25	7.93	32	9	-6.00	182.25	13.99
15	8	5.06	12.25	8.27	33	14	66.06	210.25	14.32
16	6	8.61	6.25	8.60	34	15	86.11	240.25	14.66
17	9	0.67	2.25	8.94	35	16	108.17	272.25	15.00
18	10	-0.28	0.25	9.28	36	17	132.22	306.25	15.33
					666	340	137.00	3885.00	

جدول 8-19

ب- تحديد المعاملات الفصلية (الموسمية): يتم ذلك على مرحلتين :

- نقسم القيم الحقيقية على القيم الاتجاهية ونحولها الى نسبة مائوية كما

هي واضحة في الجدول الموالي:

t	y _t	قيم الاتجاه y _t	(y _t /y _t).100	t	y _t	قيم الاتجاه y _t	(y _t /y _t).100
1	4	3.56	112.36	19	9	9.61	93.65
2	5	3.89	128.53	20	4	9.95	40.20
3	6	4.23	141.84	21	8	10.29	77.75
4	7	4.57	153.17	22	11	10.62	103.58
5	6	4.90	122.45	23	12	10.96	109.49
6	7	5.24	133.59	24	14	11.29	124.00
7	8	5.58	143.37	25	10	11.63	85.98
8	3	5.91	50.76	26	12	11.97	100.25
9	6	6.25	96.00	27	13	12.30	105.69
10	6	6.58	91.19	28	15	12.64	118.67
11	5	6.92	72.25	29	15	12.98	115.56
12	6	7.26	82.64	30	15	13.31	112.70
13	6	7.59	79.50	31	16	13.65	117.22
14	7	7.93	88.27	32	9	13.99	64.33
15	8	8.27	96.74	33	14	14.32	97.77
16	6	8.60	69.77	34	15	14.66	102.32
17	9	8.94	100.67	35	16	15.00	106.67
18	10	9.28	107.76	36	17	15.33	110.89
				666	340		مج

جدول 8-20

القيم (y_t/y_t).100 تعني القيم الحقيقية مقسومة على القيم الاتجاهية، مأخوذة كنسبة مائوية.

- نعيد ترتيب القيم ثم نوجد وسطها الحسابي كما هو واضح في العمود الأخير من الجدول 8-21، حيث نحصل على ما يسمى بالمعاملات الموسمية (الشهرية).

شهر	1	2	3	4	5	6	7	8	9	10	11	12
1992	112.3	128.5	141.8	153.2	122.5	133.6	143.4	50.8	96.0	91.2	72.2	82.4
1993	79.5	88.3	96.7	69.8	100.7	107.8	93.7	40.2	77.8	103.6	109.5	124.0
1994	86.0	100.3	105.7	118.7	115.6	112.7	117.2	64.3	97.8	102.3	106.7	110.9
معامل شهري	92.6	105.7	114.7	113.9	112.9	118.0	118.1	51.8	90.5	99.0	96.1	105.8

جدول 8-21

في شهر أبريل مثلاً يسجل إرتفاع شهري مقداره 13.9 %، بينما في شهر أكتوبر يسجل إنخفاض شهري مقداره 1 %.

ج- نحذف التغيرات الشهرية من البيانات الأولية بقسمة البيانات الحقيقية لكل شهر على المعامل الشهري وذلك ما يظهر في العمود الرابع من الجدول 8-22.

t	y _t	قيم متروعة التغيرات الشهرية	معاملات شهرية
1	4	92.6	4.3
2	5	105.7	4.7
3	6	114.7	5.2
4	7	113.9	6.1
5	6	112.9	5.3
6	7	118.0	5.9
7	8	118.1	6.8
8	3	51.8	5.8
9	6	90.5	6.6
10	6	99.0	6.1
11	5	96.1	5.2
12	6	105.8	5.7
13	6	92.6	6.5
14	7	105.7	6.6
15	8	114.7	7.0
16	6	113.9	5.3
17	9	112.9	8.0
18	10	118.0	8.5

t	y _t	قيم متروعة التغيرات الشهرية	معاملات شهرية
19	9	118.1	7.6
20	4	51.8	7.7
21	8	90.5	8.8
22	11	99.0	11.1
23	12	96.1	12.5
24	14	105.8	13.2
25	10	92.6	10.8
26	12	105.7	11.4
27	13	114.7	11.3
28	15	113.9	13.2
29	15	112.9	13.3
30	15	118.0	12.2
31	16	118.1	13.6
32	9	51.8	17.4
33	14	90.5	15.5
34	15	99.0	15.2
35	16	96.1	16.6
36	17	105.8	16.1

جدول 8-22

بإعادة ترتيب القيم نحصل في النهاية على البيانات متروعة التغيرات الشهرية وهي كما في الجدول 8-23.

شهر	1	2	3	4	5	6	7	8	9	10	11	12
1992	4.3	4.7	5.2	6.1	5.3	5.9	6.8	5.8	6.6	6.1	5.2	5.7
1993	6.5	6.6	7.0	5.3	8.0	8.5	7.6	7.7	8.8	11.1	12.5	13.2
1994	10.8	11.4	11.3	13.2	13.3	12.7	13.5	17.4	15.5	15.2	16.6	16.1

جدول 8-23

2- طريقة متوسطات الموسم : لتحديد مركبات الموسم وحذف تغيراته بهذه الطريقة تتبع الخطوات التالية:

أ- نوجد متوسطات كل موسم بتقسيم مجموع قيم كل موسم على عدد السنوات.

ب- نوجد المتوسط الموسمي العام، بتقسيم مجموع متوسطات المواسم على عدد المواسم.

ج- نوجد النسب الموسمية بقسمة متوسط كل موسم على المتوسط العام وضربه في المقدار 100 (نسبة مائوية).

د- نستبعد الاتجاه العام للظاهرة قبل حساب النسب الموسمية، وذلك بإيجاد معادلة الاتجاه العام للقيم (إعتمادا على المتوسطات السنوية)، ونحسب عن طريقها القيم الاتجاهية للظاهرة، ثم نقسم كل قيمة واقعية على القيمة الاتجاهية المقابلة لها و نضربها في 100 (نسبة مائوية)، فنحصل بذلك على نسب تدل على الظاهرة مستبعدا منها الاتجاه العام.

ج- نحذف التغيرات الموسمية بقسمة القيم الحقيقية على المعاملات الموسمية (الشهرية).

مثال 8-10: البيانات التالية خاصة بتطور أسعار الحبوب خلال الفترة 1999-2006 بالدينار للطن الواحد. المطلوب إيجاد القيم المحذوفة التغيرات الموسمية بطريقة متوسطات الموسم.

شهر سنة	1	2	3	4	5	6	7	8	9	10	11	12	معدل سري
1999	218	181	178	150	131	116	123	145	169	202	225	247	173.8
2000	242	209	199	168	149	136	142	162	188	221	242	264	193.5
2001	267	228	220	187	169	151	159	184	209	245	267	294	215.0
2002	292	249	242	211	190	173	182	205	228	264	289	317	236.8
2003	320	278	270	234	214	196	205	230	256	296	342	352	266.1
2004	353	312	298	262	241	222	125	259	292	327	354	383	285.7
2005	397	340	329	293	270	247	257	288	315	357	391	416	325.0
2006	429	377	363	323	298	280	289	319	348	393	426	460	358.8
مج	518	2174	2099	1828	1662	1521	1482	1792	2005	2305	2536	2733	
وسط	314.8	271.8	262.4	228.5	207.8	190.1	185.3	224	250.6	288.1	317	341.6	

جدول 8-24

نوجد متوسط القيم الشهرية بقسمة مجموعها على 8 (عدد السنوات).

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

نوجد المتوسط الشهري العام بقسمة مجموع المتوسطات الشهرية على 12 (عدد الأشهر)، أي:

$$(314.8+271.8+....+341.6)/12 = 256.8$$

نوجد النسب الموسمية (الشهرية) بقسمة المتوسط الشهري على المتوسط الشهري العام وضربه في 100 للحصول على نسب مائوية، كما هي واضحة في الجدول أدناه:

الشهر	1	2	3	4	5	6	7	8	9	10	11	12
متوسط شهري	314.8	271.8	262.4	228	207.8	190	185.3	224	250.6	288.1	317	341.6
نسبة شهرية	122.6	105.8	102.2	89.0	80.9	74.0	72.2	87.2	97.6	112.2	123.4	133
تغير شهري	22.6	5.8	2.2	-11	-19.1	-26	-27.8	-12.8	-2.4	12.2	23.4	33

جدول 8-25

في شهر جانفي مثلا سجل صعود موسمي يقدر بـ: 22.6 % ، بينما في شهر أفريل سجل هبوط موسمي يقدر بـ: 11%.

إذا ظهر وأن قيم الظاهرة متأثرة بإتجاه عام معين فإنه يتعين إستبعاد هذا الإتجاه قبل حساب النسب الموسمية، لأنه إذا لم تستبعد فإن التغيرات الموسمية لا تكون دقيقة، فإذا كان الإتجاه العام صعوديا مثلا، و كان التغير الموسمي صعوديا أيضا، فإن النسب الموسمية تكون أكثر مما يجب أن تكون و العكس صحيح.

لإستبعاد الإتجاه العام، نوجد معادلة الإتجاه العام للقيم الموجودة لدينا وذلك من خلال المتوسط السنوي، ونحسب القيم الإتجاهية للظاهرة، ثم نقسم كل قيمة واقعية على القيمة الإتجاهية المقابلة مأخوذة كنسبة مائوية، فنحصل على نسب تدل على الظاهرة مستبعدا منها الإتجاه العام.

بما أن عدد السنوات زوجي، و عدد الأشهر كذلك، فإن الطريقة المختصرة، تجعلنا نأخذ المتغير المستقل على أساس نصف سنوي، لإيجاد القيم الإتجاهية، ويتم ذلك عن طريق الحسابات كما هي واردة في الجداول الموالية.

و تكون معادلة الإتجاه العام على النحو:

$$y_t = a + bx_t \quad 32-8$$

حيث : y_t هي المتوسطات السنوية، يتم الحصول عليها بقسمة المجموع السنوي على 12 (انظر السطر الأخير من الجدول 8-24).
ويتم تقدير كل من : a و b بالطريقة المختصرة و ذلك بالإستعانة بالمعادلتين الطبيعيين رقم 8-10 و 8-11، أي :

$$\sum_{t=1}^n y_t = b \sum_{t=1}^n x_t + Na$$

$$\sum_{t=1}^n y_t x_t = b \sum_{t=1}^n x_t^2 + a \sum_{t=1}^n x_t$$

و بالإستعانة بالجدول التالي يتم إيجاد a و b (مع افتراض أن سنة الأساس أي السنة الأولى 1999 يكون عندها $t=1$).

t	y_t	x_t	$y_t x_t$	x_t^2
1	173.8	-7	-1216.6	49
2	193.5	-5	-967.5	25
3	215.0	-3	-645.0	9
4	236.8	-1	-236.8	1
5	266.1	1	266.1	1
6	285.7	3	857.1	9
7	325.0	5	1625.0	25
8	358.8	7	2511.6	49
مع	2054.7	0	2193.9	168

جدول 8-26

بما أن : $\sum x_t = 0$ لذلك فالقيم التقديرية لكل من : a و b تكون على النحو التالي :

$$b = \frac{\sum_{t=1}^n y_t x_t}{\sum_{t=1}^n x_t^2}$$

$$a = \frac{\sum_{t=1}^n y_t}{N}$$

$$a = \frac{2054.7}{8} = 256.84$$

ومن خلال الجدول 8-26 نجد :

$$b = \frac{2193.9}{168} = 13.06$$

و بالتالي تكون معادلة الاتجاه العام كما يلي :

$$y_t = 256.84 + 13.06x_t$$

بما أن x_t تمثل وحدات نصف سنوية إبتداء من السنة الوسيطة (نقطة الأصل)

1 جانفي 2003، فإن معامل الاتجاه العام الشهري هو :

$$\frac{a}{6} = \frac{13.06}{6} = 2.2$$

و بالتالي تحسب القيم الاتجاهية للأشهر اعتمادا على المعادلة التالية :

$$z_t = 2.2 m + 256.84 \quad 33-8$$

حيث m هو ترتيب القيم الشهرية المأخوذة في منتصف كل شهر والموضحة

في الجدول التالي :

شهر سنة	1	2	3	4	5	6	7	8	9	10	11	12
1999	-47.5	-46.5	-45.5	-44.5	-43.5	-42.5	-41.5	-40.5	-39.5	-38.5	-37.5	-36.5
2000	-35.5	-34.5	-33.5	-32.5	-31.5	-30.5	-29.5	-28.5	-27.5	-26.5	-25.5	-24.5
2001	-23.5	-22.5	-21.5	-20.5	-19.5	-18.5	-17.5	-16.5	-15.5	-14.5	-13.5	-12.5
2002	-11.5	-10.5	-9.5	-8.5	-7.5	-6.5	-5.5	-4.5	-3.5	-2.5	-1.5	-0.5
2003	0.5	1.5	2.5	3.5	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5
2004	12.5	13.5	14.5	15.5	16.5	17.5	18.5	19.5	20.5	21.5	22.5	23.5
2005	24.5	25.5	26.5	27.5	28.5	29.5	30.5	31.5	32.5	33.5	34.5	35.5
2006	36.5	37.5	38.5	39.5	40.5	41.5	42.5	43.5	44.5	45.5	46.5	47.5

جدول 8-27

و بناء على ذلك تكون القيم الاتجاهية لجميع الأشهر والسنوات كما في

الجدول 8-28 أدناه، وتم الحصول عليها بتعويض القيم كما هي في الجدول

8-27 أعلاه، على وجه الترتيب في المعادلة 8-33.

شهر سنة	1	2	3	4	5	6	7	8	9	10	11	12
1999	152.3	154.5	156.7	158.9	161.1	163.3	165.5	167.7	169.9	172.1	174.3	176.5
2000	178.7	180.9	183.1	185.3	187.5	189.7	191.9	194.1	196.3	198.5	200.7	202.9
2001	205.1	207.3	209.5	211.7	213.9	216.1	218.3	220.5	222.7	224.9	227.1	229.3
2002	231.5	233.7	235.9	238.1	240.3	242.5	244.7	246.9	249.1	251.3	253.5	255.7
2003	257.9	260.1	262.3	264.5	266.7	268.9	271.1	273.3	275.5	277.7	279.9	282.1
2004	284.3	286.5	288.7	290.9	293.1	295.3	297.5	299.7	301.9	304.1	306.3	308.5
2005	310.7	312.9	315.1	317.3	319.5	321.7	323.9	326.1	328.3	330.5	332.7	334.9
2006	337.1	339.3	341.5	343.7	345.9	348.1	350.3	352.5	354.7	356.9	359.1	361.3

جدول 8-28 economicrg

groups economicrg

economicrg.blogspot.com

الباحث الاقتصادي

235

researcher ate

للحصول على المعدلات الموسمية (الشهرية)، نقسم القيم الشهرية الأصلية على القيم الاتجاهية و نضربها في 100، فنحصل على الجدول 8-29 :

شهر سنة	1	2	3	4	5	6	7	8	9	10	11	12
1999	143.1	117.2	113.6	94.4	81.3	71.0	74.3	86.5	99.5	117.1	129.1	139.9
2000	135.4	115.5	108.7	90.7	79.5	71.7	74.0	83.5	95.8	111.3	120.6	130.1
2001	130.2	110.0	105.0	88.3	79.0	69.9	72.8	83.4	93.8	108.9	117.6	128.2
2002	126.1	106.5	102.6	88.6	79.1	71.3	74.4	83.0	91.5	105.0	114.0	124.0
2003	124.1	106.9	102.9	88.5	80.2	72.9	75.6	84.2	92.9	106.6	122.2	124.8
2004	124.2	108.9	103.2	90.1	82.2	75.2	42.0	86.4	96.7	107.5	115.6	124.1
2005	127.8	108.7	104.4	92.3	84.5	76.8	79.3	88.3	95.9	108.0	117.5	124.2
2006	127.3	111.1	106.3	94.0	86.2	80.4	82.5	90.5	98.1	110.1	118.6	127.3
معدل موسمي	129.8	110.6	105.8	90.9	81.5	73.7	71.9	85.7	95.5	109.3	119.4	127.8

جدول 8-29

السطر الأخير من الجدول أعلاه يمثل المعدلات الموسمية (الشهرية)، وتم إيجادها بقسمة مجموع النسب الشهرية على 8 (عدد السنوات).
إذا كان مجموع المعدلات الموسمية لا يساوي : 1200، فإنه يتم تصحيحها وذلك بضرب المعدل الموسمي (الشهري) في 1200 وقسمته على مجموع المعدلات الشهرية المحصل عليها سابقاً، فنجد بذلك المعدلات الموسمية المصححة.
واضح أن النتائج المتوصل إليها تحدد لنا تأثير تقلبات كل موسم على قيمة الظاهرة، وبالتالي نتمكن من الاستعداد لمواجهة الزيادة أو النقصان في الظاهرة.
يتم حذف التغيرات الشهرية بقسمة القيم الحقيقية على المعدلات الموسمية وتحويلها إلى نسبة مائوية كما جرى في المثال السابق.

تمارين

تمرين 1: لأي عنصر من عناصر السلاسل الزمنية يمكن إرجاع الحوادث التالية:

- 1- زيادة الطلب على أجهزة الإعلام الآلي منذ سنة 2000.
- 2- الحاجة الى زيادة إنتاج الحبوب نتيجة للإزدياد في عدد السكان.
- 3- ظهور فترة من الرخاء الإقتصادي.
- 4- حدوث اضطراب عام في أحد المصانع الكبرى.
- 5- زيادة إستهلاك الغاز و الكهرباء خلال شهري جانفي وفيفري.

تمرين 2: البيانات التالية تظهر تطور كميات تساقط الأمطار خلال الفترة : 1995-86 في إحدى الولايات :

السنة	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
الكمية	95	80	85	90	95	110	115	100	115	120

المطلوب:

- 1- أوجد معامل الخشونة. ماذا تستنتج؟.
- 2- عن طريق المعدل المتحرك بطول 3 أوجد سلسلة جديدة. ثم أوجد معامل خشونتها وقارنه بمعامل الخشونة المطلوب في السؤال 1.
- 3- بطريقة المتوسطات الجزئية أجب على نفس أسئلة المطلوب 2.
- 4- على معلم متعامد ارسم البيانات الأصلية و البيانات المطلوبة في السؤال 2. ماذا تستنتج؟.

تمرين 3: البيانات التالية تظهر تطور المواليد و الوفيات في الجزائر خلال الفترة 1990-2002 بآلاف الأشخاص، حسب مصادر الديوان الوطني للإحصائيات.

(المصدر / www.ons.gov.dz)

السنة	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
مواليد	775	773	799	775	776	711	654	654	607	594	589	619	617
وفيات	151	155	160	168	180	180	172	178	144	141	140	141	138

المطلوب:

- 1- أوجد معامل الخشونة لكل من المواليد و الوفيات. ماذا تستنتج؟
- 2- أوجد الاتجاه العام لكل من المواليد و الوفيات بطريقة المعدلات المتحركة ثم بطريقة المعدلات النصفية.
- 3- أوجد الاتجاه العام لكل من المواليد و الوفيات بطريقة المربعات الصغرى.
- 4- أوجد الزيادة الطبيعية للسكان عند كل سنة ثم أوجد اتجاهها العام بالطرق المطلوبة في الأسئلة السابقة.

تمرين 4: البيانات التالية تظهر تطور كل من الإنتاج الداخلي الإجمالي والواردات، خلال الفترة 1987-76 للجمهورية الجزائرية حسب أرقام الديوان الوطني للإحصائيات، بملايير الدينارات.

السنة	الواردات	إنتاج د.إ
1976	28.43	68.50
1977	37.27	81.90
1978	41.99	104.00
1979	46.47	128.50
1980	55.84	162.50
1981	65.99	191.50
السنة	الواردات	إنتاج د.إ
1982	68.32	207.60
1983	66.72	239.80
1984	68.16	259.90
1985	65.06	289.20
1986	55.79	286.50
1987	48.88	307.90

المطلوب : 1- قدم كل من الواردات و الإنتاج الداخلي الإجمالي بدلالة الزمن على محور متعامد. قارن بين المنحنيين. ماذا تستنتج؟. 2- ارسم المنحنيين المطلوبين في السؤال 1، بإعادة البيانات عن طريق معدل متحرك بطول 3 ثم بطول 4. ماذا تستنتج. 3- أوجد الدالة الزمنية للواردات و الدالة الزمنية للإنتاج الداخلي الإجمالي. 4- أوجد القيم الاتجاهية لكل من الواردات والإنتاج الداخلي الإجمالي (بدلالة الزمن) بطريقتين.

تمرين 5: البيانات التالية تظهر تطور سكان الجزائر خلال الفترة 1962 - 2003 بالملايين، حسب معطيات الديوان الوطني للإحصائيات:

السنة	1962	1963	1964	1965	1966	1967
عدد السكان	10.24	10.67	11.13	11.60	12.14	12.57
السنة	1968	1969	1970	1971	1972	1973
عدد السكان	12.95	13.35	13.75	14.18	14.16	15.07
السنة	1974	1975	1976	1977	1978	1979
عدد السكان	15.53	16.02	16.3	17.1	17.7	18.1
السنة	1980	1981	1982	1983	1984	1985
عدد السكان	18.7	19.2	19.9	20.5	21.2	21.9
السنة	1986	1987	1988	1989	1990	1991
عدد السكان	22.5	23.2	23.8	24.4	25.0	25.6
السنة	1992	1993	1994	1995	1996	1997
عدد السكان	26.3	26.9	27.4	28.0	28.6	29.0
السنة	1998	1999	2000	2001	2002	2003
عدد السكان	29.5	30.0	30.4	30.8	31.2	31.6

المطلوب: 1- أوجد النموذج الزمني المقدّر لتطور السكان.

2 - ماهو عدد السكان المتوقع لسنوات: 2010 و 2020 و 2030.

تمرين 6: البيانات التالية خاصة بتطور إنتاج الأحذية في إحدى الورشات بآلاف الأزواج. المطلوب: أوجد النموذج الزمني المقدّر للإنتاج.

السنة	1986	1987	1988	1989	1990	1991	1992	1993	1994
الإنتاج	20	30	35	40	42	35	28	18	10

تمرين 7: البيانات التالية خاصة بتطور إنتاج الحبوب في إحدى المساحات المستصلحة بعشرات القناطير:

السنة	1988	1989	1990	1991	1992	1993
الإنتاج	5	10	30	60	140	220

المطلوب: أوجد الدالة الزمنية المقدرة للإنتاج، وأوجد القيم الاتجاهية وقارنها بالقيم الحقيقية.

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018

تمرين 8: البيانات التالية تظهر الأمطار الشهرية لسنة 1992 لبعض الولايات الكبرى حسب معطيات د.و. للإحصائيات.

شهر مدن	1	2	3	4	5	6	7	8	9	10	11	12
وهران	310	180	848	134	859	237	16	9	—	264	359	178
الشلف	608	191	709	311	335	286	102	—	—	436	190	185
الجزائر	1548	410	1009	807	609	190	77	—	153	684	1397	706
سطيف	344	348	324	665	732	197	378	15	698	161	351	837
قسنطينة	528	332	495	345	973	107	173	103	206	229	1161	1928
عباية	690	805	651	115	831	149	108	6	22	364	1262	1432

المطلوب: باستخدام طريقة النسبة الى الاتجاه العام ، أوجد القيم المحذوفة التغيرات الشهرية.

تمرين 9: البيانات التالية تظهر تطور أسعار الغاز في أشهر السنة خلال الفترة 1998-2005.

شهر سنة	1	2	3	4	5	6	7	8	9	10	11	12
1998	218	181	178	150	131	116	123	145	169	202	225	247
1999	242	209	199	168	149	136	142	162	188	221	242	264
2000	267	228	220	187	169	151	159	184	209	245	267	294
2001	292	249	242	211	109	173	182	205	228	264	289	317
2002	320	278	270	234	214	196	205	230	256	296	342	352
2003	353	312	298	262	241	222	125	259	292	327	354	383
2004	387	340	329	293	270	247	257	288	315	357	391	416
2005	429	377	363	329	298	280	289	319	348	393	426	460

المطلوب: باستخدام طريقتي النسبة الى الاتجاه العام و المتوسطات الموسمية ، أوجد مركبات كل شهر و استبعد التغيرات الشهرية.

الفصل التاسع

الأرقام القياسية.

الظواهر الاقتصادية والاجتماعية تتغير من فترة لأخرى، ومن مكان لآخر، بقيم يكون ليس من السهل إدراك أهميتها عن طريق الأرقام المطلقة، وليس من السهل أيضا مقارنتها مع غيرها، فعندما نقول مثلا أن صادرات المحروقات لدولة ما إرتفعت من 35 مليار دينار سنة 1986 الى 50 مليار دينار سنة 1994، فإن قيمة التغير و هي 15 مليار دينار، لاتعطينا صورة واضحة عن أهمية هذا التطور، كما لاتعطينا أيضا صورة واضحة مقارنة بظاهرة أخرى، تغيرت خلال نفس الفترة بنفس القيمة، لذلك كان لابد من اللجوء الى التحليل النسبي للظواهر الكمية، ويتم ذلك إما عن طريق الأرقام القياسية أو عن طريق معدلات النمو عامة، و ذلك ما سنتطرق اليه في هذا الفصل.

أولا: الأرقام القياسية:

تعريفه 9-1: الرقم القياسي هو أداة لقياس التغير النسبي الحاصل في قيم أية ظاهرة أو مجموعة من الظواهر، من ظرف أول يسمى بظرف الأساس الى ظرف آخر يسمى بظرف المقارنة، سواء كان الظرف زمانيا أو مكانيا. وتكون قيمة الرقم القياسي في ظرف الأساس مساوية دائما المقدار 100. والأرقام القياسية كثيرة الإستخدام في دراسة تطوّر الظواهر الاقتصادية والاجتماعية، كالإنتاج، الإستهلاك، الصادرات، الواردات... الخ، غير أنها أكثر إستخداما في دراسة تطور الأسعار والنفقات الإستهلاكية، لذلك سوف نجعل كل التعاريف والقوانين الموالية منطبقة عليها.

هناك ثلاثة أصناف من الطرق لإيجاد الأرقام القياسية هي :

1- الأرقام القياسية البسيطة. 2- الأرقام القياسية المرجحة. 3- الأرقام القياسية ذات الأساس المتحرك.

1-الأرقام القياسية البسيطة : تضم مجموعة من الطرق هي: طريقة المناسيب البسيطة، طريقة الوسط الحسابي للمناسيب البسيطة، طريقة الوسط

فقط للاستعمال الشخصي economicrg.blogspot.com بوابة الباحث الاقتصادي © 2018
الهندسي للمناسيب البسيطة و الطريقة التجميعية المرجحة، و ستطرق لكل
واحدة فيما يلي:

1- طريقة المناسيب البسيطة :

تعريف 9-2: اذا كان سعر مادة ما في الفترة 0 (فترة الأساس) هو P_0 ،
وأصبح سعرها في الفترة 1 (فترة المقارنة): P_1 ، فإن الرقم القياسي بطريقة
المناسيب البسيطة يعرف كما يلي:

$$I = \frac{P_1}{P_0} \times 100 \quad 1-9$$

يعني هذا أنه إذا كان سعر المادة 100 وحدة نقدية بأسعار فترة الأساس، فإنه
يصبح يساوي I بأسعار فترة المقارنة.

مثال 9-1: إذا كان سعر الكيلوغرام الواحد من السكر سنة : 2002 هو : 35
دينار، وأصبح سعره سنة : 2003 40 دينار، أوجد الرقم القياسي لتطور سعر
السكر بطريقة المناسيب البسيطة.

الجواب:

$$I = \frac{P_{2003}}{P_{2002}} \times 100 = \frac{40}{35} \times 100 = 114.28 \% \quad 2-9$$

يعني هذا أنه إذا كان سعر السكر يساوي 100 دينار بأسعار سنة 2002، فإنه
ارتفع ليصبح 114.28 دينار بأسعار سنة 2003.

تستخدم هذه الطريقة، في دراسة تطور قيمة ظاهرة واحدة فقط، غير أنه عمليا،
وفي الكثير من الأحيان يتطلب الأمر، دراسة تطور أسعار عدة مواد، وفي مثل
هذه الحالة غالبا ما يتم استخدام الوسط الحسابي أو الوسط الهندسي للمناسيب
البسيطة.

ب- طريقة الوسط الحسابي للمناسيب البسيطة:

تعريف 9-3: إذا كانت لدينا : $P_{0,1} , P_{0,2} , P_{0,3} \dots P_{0,n}$ ، أسعار
فترة الأساس للمواد : 1، 2، 3 . . . N على التوالي، و : $P_{1,1} , P_{1,2} , \dots$

فإن الوسط الحسابي للمناسيب البسيطة يعرف بالمعادلة التالية:

$$I = \frac{\sum_{i=1}^n \left[\frac{P_{1,i}}{P_{0,i}} \right]}{N}$$

3-9

ج- طريقة الوسط الهندسي للمناسيب البسيطة:

تعريف 9-4: إذا كانت لدينا: $P_{0,1}, P_{0,2}, P_{0,3}, \dots, P_{0,n}$

هي أسعار سنة الأساس للمواد: 1، 2، 3، ... N على التوالي

و $P_{1,1}, P_{1,2}, P_{1,3}, \dots, P_{1,n}$

هي أسعار سنة المقارنة لنفس المواد، حيث N هو عدد المواد، فإن الوسط الهندسي للمناسيب البسيطة يعرف، بأنه الجذر النوني لجداءات قيم المناسيب البسيطة، أي:

$$I = \sqrt[n]{\prod_{i=1}^n \left[\frac{P_{1,i}}{P_{0,i}} \right]} \times 100$$

4-9

مثال 9-2: البيانات التالية تظهر تطور أسعار مجموعة من المواد الإستهلاكية بين سنتي: 1992 و 1993، بالدينارات.

المطلوب: أوجد الرقم القياسي بطريقتي الوسط الحسابي والوسط الهندسي للمناسيب البسيطة.

المادة/السعر	P ₁₉₉₂	P ₁₉₉₃
الخبز (كلغ)	1.5	2.5
الحليب (ل)	3.0	4.0
الزيت (5ل)	140.0	150.0
القهوة (كلغ)	30.0	34.0
السكر (كلغ)	10.0	15.0
الطماطم (كلغ)	25.0	50.0
الصابون (قطعة)	10.0	12.5

الإجابة: لإيجاد الوسط الحسابي و الوسط الهندسي للمناسيب البسيطة، يتطلب الأمر أولاً إيجاد المناسيب البسيطة باستخدام المعادلة رقم 9-1، ثم تطبيق المعادلتين، 9-3 و 9-4 على التوالي، وذلك بمساعدة الجدول التالي:

i	المادة/السعر	P _{0,i}	P _{1,i}	$\frac{P_{1,i}}{P_{0,i}} \times 100$
1	الخبز (كغ)	1.5	2.5	166.67
2	الحليب (ل)	3.0	4.0	133.33
3	الزيت (5ل)	140.0	150.0	107.14
4	القهوة (كغ)	30.0	34.0	113.33
5	السكر (كغ)	10.0	15.0	150.00
6	الطماطم (كغ)	25.0	50.0	200.00
7	الصابون (قطعة)	10.0	12.5	125.00
مج		219.5	268	995.47

جدول 9-2

• **الوسط الحسابي للمناسيب البسيطة:** بتطبيق المعادلة 9-3 نجد :

$$I = \frac{995.47}{7} \times 100 = 142.21 \%$$

يعني هذا أنه إذا كانت تكاليف شراء تشكيلة المواد المشار إليها في الجدول هي 100 دينار بأسعار سنة 1992، فإن نفس تشكيلة المواد أصبحت تكلف 142.21 دينار بأسعار سنة 1993.

• **الوسط الهندسي للمناسيب البسيطة:** بتطبيق المعادلة 9-4 نجد :

$$I = \sqrt[7]{10.118} \times 100 = 139.18 \%$$

مدلول ذلك أنه إذا كانت تكاليف شراء تشكيلة المواد المشار إليها في الجدول هي 100 دينار بأسعار سنة 1992، فإن نفس تشكيلة المواد أصبحت تكلف 139.18 دينار بأسعار سنة 1993.

د- الطريقة التجميعية البسيطة: تعرف كما يلي :

تعريف 9-4: إذا كانت لدينا : $P_{0,1}, P_{0,2}, P_{0,3}, \dots, P_{0,n}$ أسعار فترة الأساس للمواد : 1، 2، 3، ...، N على التوالي

الباحث الاقتصادي

و $P_{1,1} , P_{1,2} , P_{1,3} \dots P_{1,n}$

هي أسعار فترة المقارنة لنفس المواد، حيث N : عدد المواد، فإن الرقم القياسي بالطريقة التجميعية البسيطة، يعطى كما يلي:

$$I = \frac{\sum_{i=1}^n P_{1,i}}{\sum_{i=1}^n P_{0,i}} \times 100 \quad 5-9$$

مثال 9-3 ، أوجد الرقم القياسي بالطريقة التجميعية البسيطة لبيانات المثال 9-2. بتطبيق المعادلة رقم 9-5 نجد:

$$I = \frac{268}{219.5} \times 100 = 122 \%$$

بتعميق النظر في مجاميع الأسعار في جدول المثال 9-2 وفي أسعار كل مادة على حدة، نجد أن أسعار كل من الزيت و البن هيمن على مجموع الأسعار، بينما أسعار بقية المواد لا تشكل سوى جزءا ضئيلا من مجموع الأسعار، و نتيجة لهذا فإن الرقم القياسي المحصل عليه بهذه الطريقة يتأثر تأثيرا كبيرا بالمواد المرتفعة الثمن، وفي الواقع العملي وحسب النظرية الإقتصادية الجزئية، فإن المواد المرتفعة الثمن، هي أقل طلبا وأقل أهمية بالنسبة للمستهلك، لذلك فإن هذه الطريقة (الأرقام القياسية البسيطة)، معيبة، وهي أقل استخداما، ويكون من الضروري ادخال الكميات المستهلكة من كل سلعة كأوزان أي ترجيح الأسعار بالكميات، لإيجاد رقم قياسي أفضل، يعتمد على مجموع النفقات على السلع الاستهلاكية، وليس على الأسعار وحدها، وهذا ما نستعرضه في الصنف الثاني من الأرقام القياسية.

2-الأرقام القياسية المرجحة: في هذه الطرق يتم أخذ الكميات المستهلكة من كل مادة كأوزان، بحيث تنتقل من الرقم القياسي التجميعي البسيط للأسعار الى الرقم القياسي التجميعي للنفقات على المواد، وذلك لتفادي عيب الطرق السابقة، بحيث يتم حصر المواد المستهلكة وأسعارها والكميات

المستهلكة من كل مادة، لعائلة استهلاكية، أو لمجتمع استهلاكي يتكون من عدد محدد من الأفراد، و يتم استخدام إحدى الطرق التالية:

1- الطريقة التجميعية المرجحة:

تعريف 5-9: إذا كانت لدينا : $P_{0,1} , P_{0,2} , P_{0,3} \dots P_{0,n}$

أسعار فترة الأساس للمواد : 1، 2، 3 . . . N على التوالي

و $P_{1,1} , P_{1,2} , P_{1,3} \dots P_{1,n}$

هي أسعار فترة المقارنة لنفس المواد، وأن الكميات المستهلكة من كل مادة في

الفترتين هي : $Q_1 , Q_2 , Q_3 \dots Q_n$ على التوالي، حيث N : عدد المواد، فإن

الرقم القياسي بالطريقة التجميعية المرجحة يعطى كما يلي:

$$I = \frac{\sum_{i=1}^n P_{1,i} Q_i}{\sum_{i=1}^n P_{0,i} Q_i} \times 100$$

6-9

مثال 4-9: الجدول التالي يظهر تطور أسعار مجموعة من المواد بالدينار، وكذا

الكميات المستهلكة منها، لعائلة تتكون من 6 أفراد. المطلوب إيجاد الرقم القياسي

بالطريقة التجميعية المرجحة.

i	المادة/السعر	$P_{0,i}$	$P_{1,i}$	Q_i
1	الخبز (كغ)	1.5	2.5	20
2	الحليب (ل)	3.0	4.0	30
3	الزيت (5ل)	140.0	150.0	1
4	القهوة (كغ)	30.0	34.0	1
5	السكر (كغ)	10.0	15.0	4
6	الطماطم (كغ)	25.0	50.0	2
7	الصابون (قطعة)	10.0	12.5	4

جدول 3-9

حيث : Q_i : الكميات المستهلكة من المادة i .

الإجابة: لإيجاد الرقم القياسي المطلوب، نوجد مجموع أسعار سنة الأساس مضروبة في الكميات، ومجموع أسعار سنة المقارنة مضروبة في الكميات، كما هو واضح في الجدول 4-9 أدناه .

i	المادة/السعر	P _{0,i}	P _{1,i}	Q _i	P _{0,i} .Q _i	P _{1,i} .Q _i
1	الخبز (كغ)	1.5	2.5	20	30	50
2	الحليب (ل)	3.0	4.0	30	90	120
3	الزيت (5ل)	140.0	150.0	1	140	150
4	القهوة (كغ)	30.0	34.0	1	30	34
5	السكر (كغ)	10.0	15.0	4	40	60
6	الطماطم (كغ)	25.0	50.0	2	50	100
7	الصابون (قطعة)	10.0	12.5	4	40	50
					420	564

جدول 4-9

بتطبيق المعادلة 6-9 نجد :

$$I = \frac{564}{420} \times 100 = 134.29 \%$$

يعني هذا أنه إذا كانت نفقات إستهلاك تشكيلة المواد المشار إليها هي 100 دينار بأسعار سنة 1992، فإن نفس تشكيلة المواد أصبحت تكلف حوالي 134 دينار بأسعار سنة 1993.

لو قارنا نتائج الطريقة التجميعية البسيطة، مع نتائج الطريقة التجميعية المرجحة، لوجدنا أن هناك فرق ناتج عن ترجيح الأثمان بالكميات المستهلكة في الفترتين.

إن هذه الطريقة تفترض ثبات الكميات المستهلكة في الفترتين، أي أن المستهلك يبقى يستهلك نفس الكميات مهما تغيرت الأسعار، غير أن واقع سلوك المستهلك غير ذلك، فالكميات التي يستهلكها في سنة الأساس عندما تكون الأسعار عند مستوى معين، هي غير الكميات التي يستهلكها من كل مادة عندما تتغير الأسعار في فترة لاحقة، لذلك فهناك طرق أخرى تأخذ ذلك بالحسبان، لأجل إيجاد أنسب رقم قياسي.

بجـ - طريقة لاسبير (Laspeyres): وفيها يتم إستخدام كميات فترة الأساس كأوزان، وتعطى بالمعادلة التالية:

$$I = \frac{\sum_{i=1}^n P_{1,i} \times Q_{0,i}}{\sum_{i=1}^n P_{0,i} \times Q_{0,i}} \times 100 \quad 7-9$$

حيث : $P_{1,i}$: سعر فترة المقارنة للمادة i . $P_{0,i}$: سعر فترة الأساس للمادة i .

$Q_{0,i}$: الكميات المستهلكة من المادة i في فترة الأساس.

جـ - طريقة باش (Paasche): وفيها يتم استخدام كميات فترة المقارنة كأوزان، وتعطى بالمعادلة التالية :

$$I = \frac{\sum_{i=1}^n P_{1,i} \times Q_{1,i}}{\sum_{i=1}^n P_{0,i} \times Q_{1,i}} \times 100 \quad 8-9$$

حيث : $Q_{1,i}$: كميات فترة المقارنة للمادة i .

دـ - طريقة فيشر (Fisher): وهي تقوم على أساس الجمع بين طريقتي لاسبير وباش، اذ يتم ايجاد الرقم القياسي عن طريق الوسط الهندسي لرقمي لاسبير وباش، حيث يتم الحصول على رقم تتوفر فيه جميع الصفات المطلوبة في الرقم القياسي الصحيح، لذلك يسمى هذا الرقم بالرقم القياسي الأمثل، ويعطى عن طريق المعادلة التالية:

$$I = \sqrt{\frac{\sum_{i=1}^n P_{1,i} \times Q_{0,i}}{\sum_{i=1}^n P_{0,i} \times Q_{0,i}} \times \frac{\sum_{i=1}^n P_{1,i} \times Q_{1,i}}{\sum_{i=1}^n P_{0,i} \times Q_{1,i}}} \times 100 \quad 9-9$$

مثال 9-5: من بيانات الجدول التالي، أوجد الأرقام القياسية التالية:

1- رقم باش. 2- رقم لاسبيرس. 3- الرقم الأمثل.

Q _{1,i}	Q _{0,i}	P _{1,i}	P _{0,i}	المادة/السعر	i
18	20	2.5	1.5	الحبز (كلف)	1
25	30	4.0	3.0	الحليب (ل)	2
0.8	1	150.0	140.0	الريت (ل)	3
1	1	34.0	30.0	القهوة (كلف)	4
3.5	4	15.0	10.0	السكر (كلف)	5
1.5	2	50.0	25.0	الطماطم (كلف)	6
4	4	12.5	10.0	الصابون (قطعة)	7
		268	219.5		مج

جدول 9-5

حيث: P_{0,i}: سعر سنة الأساس للمادة i. Q_{0,i}: كمية سنة الأساس للمادة i.

P_{1,i}: سعر سنة المقارنة للمادة i. Q_{1,i}: كمية سنة المقارنة للمادة i.

لإيجاد الأرقام القياسية المطلوبة نستخدم المعادلات أعلاه و بمساعدة الجدول:

Q _{1,i} P _{0,i}	P _{1,i} Q _{0,i}	P _{1,i} Q _{1,i}	P _{0,i} Q _{0,i}	Q _{1,i}	Q _{0,i}	P _{1,i}	P _{0,i}	مادة/سعر
27	50	45	30	18	20	2.5	1.5	حز
75	120	100	90	25	30	4.0	3.0	حليب
112	150	120	140	0.8	1	150.0	140	ريت
30	34	34	30	1	1	34.0	30.0	قهوة
35	60	52.5	40	3.5	4	15.0	10.0	سكر
37.5	100	75	50	1.5	2	50.0	25.0	طماطم
40	50	50	40	4	4	12.5	10.0	صابون
356.5	564	476.5	420					

نجد:

* رقم لاسبيرس : باستخدام المعادلة رقم 7-9 نجد :

$$I = (564/420) \cdot 100 = 134.29$$

* رقم باش : باستخدام المعادلة رقم 8-9 نجد :

$$I = (476.5/356.5) \cdot 100 = 133.66$$

* رقم فيشر : باستخدام المعادلة رقم 9-9 نجد :

$$I = \sqrt{134.29 \times 133.66} = 133.97$$

و تفسر هذه الأرقام بنفس التفسير السابقة .

3- الأرقام القياسية ذات الأساس المتحرك: إذا كانت المدة

بين فترة المقارنة وفترة الأساس طويلة، فإن الرقم القياسي قد لا يعبر تعبيرا صحيحا عن تطور الظاهرة، لذلك فإنه يلجأ أحيانا الى استخدام أساس متحرك على طول سلسلة البيانات، بحيث نوجد الرقم القياسي لكل فترة بالنسبة للفترة السابقة لها بطريقة الوسط الحسابي للمناسيب البسيطة، و ذلك كما يلي :

$$I_{t+1} = \sum_{i=1}^n \frac{P_{t+1,i}}{P_{t,i}} \times \frac{1}{N} \quad 10-9$$

حيث : $P_{t+1,i}$: سعر المقارنة في الفترة : $t+1$ ، للمادة i .

$p_{t,i}$: سعر الأساس في الفترة t السابقة للفترة $t+1$ ، للمادة i .

وعند الحصول على جميع الأرقام القياسية بهذه الطريقة (معادلة 9-10)، نضربها في بعضها البعض، ونضرب النتيجة في 100 فنحصل بذلك على رقم قياسي يعكس تطور الظاهرة بين فترة المقارنة وهي آخر فترة ضمن سلسلة البيانات وفترة الأساس وهي أول فترة ضمن سلسلة البيانات، ويسمى الرقم المحصل عليه بالرقم القياسي ذي الأساس المتحرك.

مثال 9-6: البيانات التاية تظهر تطور أسعار 3 مواد أساسية خلال الفترة 1989-1994 بالدينارات، المطلوب إيجاد الرقم القياسي للتطور بين سنتي 1989 و 1994 باستخدام طريقة المعدلات المتحركة.

سنة \ سلعة	1989	1990	1991	1992	1993	1994
أ	10	12	12	15	20	30
ب	15	15	20	20	21	21
ج	30	30	30	30	32	32

جدول 9-7

بتطبيق المعادلة رقم 9-10 نجد المعدلات المتحركة التالية:

$$\left(\frac{12}{10} + \frac{15}{15} + \frac{30}{30}\right) \times \frac{1}{3} = 1.07 \quad \text{1990 بالنسبة لـ: 1989}$$

$$\left(\frac{12}{12} + \frac{20}{15} + \frac{30}{30}\right) \times \frac{1}{3} = 1.11 \quad \text{1991 بالنسبة لـ: 1990}$$

$$\left(\frac{15}{12} + \frac{20}{20} + \frac{30}{30}\right) \times \frac{1}{3} = 1.08 \quad \text{1992 بالنسبة لـ: 1991}$$

$$\left(\frac{20}{15} + \frac{21}{20} + \frac{32}{30}\right) \times \frac{1}{3} = 1.15 \quad \text{1993 بالنسبة لـ: 1992}$$

$$\left(\frac{30}{20} + \frac{21}{21} + \frac{32}{32}\right) \times \frac{1}{3} = 1.17 \quad \text{1994 بالنسبة لـ: 1993}$$

وعليه يكون الرقم القياسي لأسعار سنة 1994 بالنسبة لسنة 1989 بطريقة الأساس المتحرك هو :

$$I = (1.07 \times 1.11 \times 1.08 \times 1.15 \times 1.17) 100 = 172.59$$

هذه هي أهم أنواع الأرقام القياسية المستخدمة في الحياة العملية، و كما رأينا فمهما كان نوعها تبقى مقياسا نسبيا لتطور الظواهر تتأثر كثيرا بطبيعة فترة الأساس و بمدى موضوعية الباحث الإحصائي في البحث عن أرقام قياسية تعكس بأكبر قدر ممكن طبيعة تطور هذه الظواهر.

4- اختيار فترة الأساس : من المشكلات التي يصادفها الإحصائي عند إعداد الأرقام القياسية هي مشكلة اختيار سنة الأساس، والإحصائي الماهر هو الذي يجيد ذلك، بحيث يجب أن يحرص على أن تكون فترة الأساس فترة عادية خالية من حالات الشذوذ، أو الحالات الطارئة، وذلك لتعكس بحسب طبيعة تطور الظاهرة، وإذا حدث وأن

تعذر إيجاد مثل هذه الفترة، فإنه لابد من إجتنا ب إتخاذ فترة واحدة كأساس، وذلك، إما بجعل متوسط قيم مجموعة من الفترات كأساس، أو إستخدام طريقة المعدلات المتحركة.

5- خصائص الأرقام القياسية الموضوعية : حتى يكون الرقم القياسي في أحسن صيغة تجعله أكثر موضوعية يجب أن يتصف بمجموعة من الخصائص منها ما يلي:

أ- خاصية المطابقة : إذا كانت قيمة الفترة 0 (الأساس) تساوي قيمة الفترة 1 (المقارنة) فإن الرقم القياسي للفترة 1 بالنسبة للفترة 0 يساوي الرقم القياسي للفترة 0 بالنسبة للفترة 1 ويساوي 100 % .

ب- خاصية الجداء : جذر جداء الرقم القياسي للفترة 1 بالنسبة للفترة 0 و الرقم القياسي للفترة 0 بالنسبة للفترة 1 يساوي الى : 100 % . أي أن حاصل جذر جداء الرقمين القياسين المتبادلين يساوي 100 % أو 1 إذا كانا غير مضروبين في 100 .

مثال 9-7: أثبت الخاصية ب حسايبا إنطلاقا من بيانات المثال 9-1 .

الجواب : الرقم القياسي البسيط لسنة 2003 بالنسبة لسنة 2002 هو :

$$I_1 = (40/35) \cdot 100 = 114.29$$

الرقم القياسي البسيط لسنة 2002 بالنسبة لسنة 2003 هو :

$$I_2 = (35/40) \cdot 100 = 87.5$$

$$\sqrt{I_1 \times I_2} = \sqrt{114.29 \times 87.5} = 100 \%$$

فالخاصية إذن صحيحة.

ج- خاصية التعدي : إذا كانت لدينا الفترات 0، 1، 2 فإن الرقم القياسي للفترة 2 بالنسبة للفترة 0 يساوي جداء الرقم القياسي للفترة 2 بالنسبة للفترة 1 في الرقم القياسي للفترة 1 بالنسبة للفترة 0، أي :

$$I_{2/0} = I_{2/1} \cdot I_{1/0} \quad 11-9$$

مثال 9-9: إذا كانت لدينا أسعار مادة ما خلال 3 فترات هي :

الفترة 0 : 10 ، الفترة 1 : 15 ، الفترة 2 : 20

الباحث الإقتصادي

تأكد من صحة الخاصية ج .

الجواب: بتطبيق المعادلة 9-11 نجد:

$$I_{2/0} = (20/15) \cdot (15/10) = 20/10 = 200\%$$

و هي نفس القيمة المحصل عليها فيما لو استخدمنا الطريقة المباشرة، و بالتالي فإن الخاصية صحيحة.

د- خاصية التجانس: وهي تعني أن الرقم القياسي قيمة نسبية مستقلة عن وحدات القياس.

هـ- خاصية الانعكاس في الأساس: الرقم القياسي للفترة 1 بالنسبة للفترة 0 يساوي مقلوب الرقم القياسي للفترة 0 بالنسبة للفترة 1، إذا كانا مأخوذين كنسبة واحدة وليس مائوية، أي :

$$I_1 = \frac{1}{I_2}$$

12-9

تستخدم هذه الخاصية لإختبار الانعكاس في الأساس للرقم القياسي.

مثال 9-8: أثبت الخاصية هـ انطلاقاً من بيانات المثال 9-1.

الرقم القياسي البسيط لسنة 2003 بالنسبة لسنة 2002 هو :

$$I_1 = (40/35) = 1.1429 \text{ أي } I_1 = 114.29\%$$

الرقم القياسي البسيط لسنة 2002 بالنسبة لسنة 2003 هو :

$$I_2 = (35/40) = 0.875 \text{ أي } I_2 = 87.5\%$$

و منه نجد أن :

$$I_1 = \frac{1}{I_2} = \frac{1}{0.66} = 1.5 \text{ أي } I_1 = 114.29\%$$

و هي نفس القيمة المحصل عليها أعلاه، و بالتالي الخاصية صحيحة.

و- خاصية الانعكاس في العامل: تعني هذه الخاصية أن الرقم القياسي للأسعار مضروباً في الرقم القياسي للكميات يساوي الرقم القياسي للقيمة. تستعمل هذه الخاصية فيما يسمى بإختبار الانعكاس في العامل، ولا يوجد سوى

رقم واحد فقط يحقق هذه الخاصية وهو الرقم القياسي الأمثل، ويمكن إثبات ذلك كما يلي :

الرقم القياسي الأمثل للأسعار (مرجحا بالكميات) هو :

$$I_q = \sqrt{\frac{\sum_{i=1}^n P_{1,i} \times Q_{0,i}}{\sum_{i=1}^n P_{0,i} \times Q_{0,i}} \times \frac{\sum_{i=1}^n P_{1,i} \times Q_{1,i}}{\sum_{i=1}^n P_{0,i} \times Q_{1,i}}}$$

الرقم القياسي الأمثل للكميات (مرجحا بالأسعار) هو :

$$I_p = \sqrt{\frac{\sum_{i=1}^n Q_{1,i} \times P_{0,i}}{\sum_{i=1}^n Q_{0,i} \times P_{0,i}} \times \frac{\sum_{i=1}^n Q_{1,i} \times P_{1,i}}{\sum_{i=1}^n Q_{0,i} \times P_{1,i}}}$$

ومنه يكون الجداء : $I_q \cdot I_p$ يساوي الرقم القياسي للقيم أي :

$$I_q \times I_p = \frac{\sum_{i=1}^n P_{1,i} \cdot Q_{1,i}}{\sum_{i=1}^n P_{0,i} \cdot Q_{0,i}}$$

ينبغي ضرب النتيجة في 100 لتكون نسبة مائوية.

إن أفضل رقم قياسي هو ذلك الذي يحقق مجمل هذه الخصائص، ويقتصر الكثير من الإحصائيين على أن الرقم القياسي الأفضل هو الذي يحقق خاصية الانعكاس في الأساس وخاصية الانعكاس في العامل، وذلك ما يطلق عليه عادة اختبار الأرقام القياسية.

معدلات النمو: معدل نمو ظاهرة ما هو النسبة المئوية لتغيرها من فترة لأخرى ، و يمكن إيجاد هذا المعدل إما عن طريق معدل النمو البسيط أو معدل النمو السنوي المتوسط.

1- معدل النمو البسيط

تعريفه 5-9: إذا كانت قيمة ظاهرة ما في فترة الأساس هي X_0 ، وأصبحت في فترة المقارنة X_1 ، فإن معدل نمو هذه الظاهرة خلال الفترة يعطى كما يلي :

$$T = \frac{X_1 - X_0}{X_0} \times 100 \quad 13-9$$

و يمكن كتابة العبارة 13-9 على النحو:

$$T = \frac{X_1}{X_0} \times 100 - 100 \quad 14-9$$

$$T = \left(\frac{X_1}{X_0} - 1 \right) \times 100 \quad \text{أو:}$$

معلوم أن الطرف الأيمن من المعادلة 14-9، هو الرقم القياسي بطريقة المناسب البسيطة كما هو معرف في المعادلة 1-9، لذلك يمكن كتابة معدل النمو كذلك كما يلي:

$$T = I - 100 \quad 15-9$$

أي الرقم القياسي منقوصاً منه 100.

يمكن إيجاد معدل النمو اعتماداً على أي من طرق الأرقام القياسية، وذلك بطرح المقدار 100 من الرقم القياسي، فنجد بذلك معدل النمو باستخدام الطريقة التجميعية البسيطة أو معدل النمو عن طريق الوسط الحسابي أو الوسط الهندسي للمناسيب البسيطة... الخ.

2- معدل النمو السنوي المتوسط

تعريفه 6-9: إذا كانت لدينا مجموعة من البيانات، واخترنا منها : X_1 قيمة سنة المقارنة و X_0 قيمة سنة الأساس، فإن معدل النمو السنوي المتوسط بين الفترتين 0 و 1 لهذه البيانات يعطى كما يلي:

$$T = \frac{\text{Log } X_1 - \text{Log } X_0}{N \cdot \text{Log } e} \times 100 \quad 16-9$$

حيث N : عدد السنوات. e : أساس اللوغاريتم النيبيري أي : $e=2.718$ ، $\text{Log } e=0.43429$

و يمكن استخدام العبارة التالية أيضا لإيجاد معدل النمو السنوي المتوسط :

$$T = \left[\left[\frac{X_1}{X_0} \right]^{\frac{1}{n}} - 1 \right] \times 100$$

17-9

مثال 9-9: البيانات التالية تظهر تطور إنتاج الحبوب في إحدى المزارع .

السنة	1989	1990	1991	1992	1993	1994	1995
الكمية	95	80	110	115	100	115	120

جدول 9-9

المطلوب : أوجد معدل النمو السنوي المتوسط للإنتاج.

الإجابة : $X_0=95$, $X_1=120$

$$T = \frac{\text{Log}120 - \text{Log}95}{7\text{Log } e} \times 100 = 3.34 \%$$

ويعني هذا أن معدل النمو السنوي المتوسط للإنتاج بين سنتي 1989 و 1995 هو 3.34 % .

تمارين

تمرين 1: ما هي الإحتياطات النظرية التي يجب على الإحصائي الإلمام بها لإعداد أرقام قياسية موضوعية؟

تمرين 2: البيانات التالية تظهر تطور عدد سكان الجزائر خلال السنوات 1980-2003 بالملايين حسب معطيات الديوان الوطني للإحصائيات:

السنة	1980	1981	1982	1983	1984	1985
عدد السكان	18.7	19.2	19.9	20.5	21.2	21.9
السنة	1986	1987	1988	1989	1990	1991
عدد السكان	22.5	23.2	23.8	24.4	25.0	25.6
السنة	1992	1993	1994	1995	1996	1997
عدد السكان	26.3	26.9	27.4	28.0	28.6	29.0
السنة	1998	1999	2000	2001	2002	2003
عدد السكان	29.5	30.0	30.4	30.8	31.2	31.6

المطلوب :

1- أوجد الأرقام القياسية البسيطة لتطور السكان و اشرحها وذلك بإعتماد:

أ- السنة السابقة كسنة أساس.

ب- سنة 1980 سنة أساس

ج- سنة 1990 سنة أساس

د- سنة 2000 سنة أساس.

2- أوجد معدل النمو للحالات أ ، ب ، ج ، د

3- أوجد معدل النمو السنوي المتوسط لتطور السكان.

تمرين 3: البيانات التالية تظهر تطور أسعار سلة من المواد الغذائية في الجزائر خلال الفترة 1998-2000 بالدينار.

المصدر: الجزائر بالأرقام. العدد 31. الديوان الوطني للإحصائيات www.ons.dz

المادة	1998	1999	2000
حليب (ل)	20	20	20
زبدة (كغ)	292.45	320.42	332.52
زيت المائدة (ل)	395.85	357.12	339.63
فاصوليا (كغ)	67.08	79.38	81.39
بطاطس (كغ)	26.92	26.37	22.05
بصل (كغ)	21.21	26.05	29.52
طماطم (كغ)	36.24	33.48	33.20
برتقال (كغ)	53.30	56.20	62.18
تمر معلب	127.37	125.88	114.88
سكر قطع (كغ)	58.17	59.14	59.69
سكر مبشور (كغ)	40.33	36.90	37.38
خبز (250 غ)	8.39	8.39	8.37
دقيق (50 كغ)	1555.88	1555.88	1558.38
كسكس (500 غ)	34.71	35.00	35.00
عجائن غذائية (كغ)	70.00	73.12	73.33
لحم بقر ب ع (كغ)	594.21	631.87	636.92
لحم خروف (كغ)	483.84	498.95	466.40
لحم دجاج (كغ)	158.22	150.87	152.04

المطلوب:

- 1- أوجد الرقم القياسي البسيط لكل مادة مرة باعتماد سنة 1998 كسنة أساس و مرة أخرى باعتماد سنة 1999 كسنة أساس. علق على النتيجة في كل حالة.
- 2- أوجد الرقم القياسي بطريقة الوسط الحسابي للمناسيب البسيطة عند كل سنة، علق على النتيجة.
- 3- أوجد الرقم القياسي بطريقة الوسط الهندسي للمناسيب البسيطة عند كل سنة، علق على النتيجة.

4- أوجد الرقم القياسي بالطريقة التجميعية البسيطة عند كل سنة، علق على النتيجة.

5- قدر الإستهلاك الشهري لعائلتك من كل مادة من المواد المذكورة في الجدول واحسب الرقم القياسي المرجح بالكميات بالإعتماد على أسعار سنة 2000.

6- أوجد معدل النمو في كل حالة من الحالات المطلوبة في الأسئلة السابقة.

تمرين 4: البيانات التالية خاصة بتطور متوسط أسعار مجموعة من المواد بالأسعار الجارية في الجزائر ، حسب أرقام الديوان الوطني للإحصائيات. (بالدينار الجزائري)

المادة/سنة	1990	1991	1992	1996
خبز (250 غ)	0.83	0.83	1.39	7.5
أرز (كغ)	7.60	9.16	14.22	47.0
لحم خروف (كغ)	142.15	181.02	211.93	370.0
لحم دجاج (كغ)	43.34	54.37	63.55	120.0
حليب (ل)	1.75	1.84	3.17	38.0
زبدة (كغ)	35.02	76.34	128.25	230.0
زيت (5ل)	24.92	27.49	75.66	320.0
عدس (كغ)	8.15	10.12	15.40	70.0
فاصولياء (كغ)	7.89	10.13	14.89	75.0
بطاطا (كغ)	7.74	9.39	8.14	30.0
سكر (كغ)	3.69	5.29	14.71	42.0
قهوة (250 غ)	25.54	26.53	30.21	120.0

المطلوب: 1- أوجد جميع الأرقام القياسية الممكنة وقارن بينها ، وذلك بإستعمال سنة 1990 كأساس ، ثم سنة 1992 كأساس .

2- أوجد الأرقام القياسية: باش ، لاسبيرس ، فيشر ، وذلك حسب الكميات المستهلكة أسبوعيا في أسرتك خلال السنوات 1990 ، 1991 ، 1992 ، 1996 ، وذلك بإستعمال مرة سنة 1990 كأساس ، وأخرى سنة 1992 كأساس.

تمرين 5: البيانات التالية تظهر تطور أسعار المواد الأولية و الكميات المستخدمة لإنتاج الأحذية في إحدى الورشات خلال الفترة 1990-1995:

المواد	الأسعار		الكميات	
	1995	1990	1995	1990
جلود طبيعية	20	25	10	12
جلود بلاستيكية	5	15	12	8
صفائح بلاستيكية	10	12	5	5
غراء	5	8	15	10
طلاء	4	4	20	20

المطلوب : 1- أوجد و اشرح الأرقام القياسية باستعمال:

- أ- طريقة لاسبير. ب- طريقة باش. ج- طريقة فيشر.
 2- أوجد الرقم القياسي المرجح بالوسط الحسابي للكميات.
 3- استنتج معدلات النمو من الطرق : ا ، ب ، ج و فسرهما.

تمرين 6: البيانات التالية خاصة بتطور أسعار مجموعة من السلع الإستهلاكية الأساسية بالدينارات الجزائرية .

مادة/سنة	1989	1990	1991	1992	1993	1994	1995
حليب	1.5	2.0	2.5	3.0	5.5	8.5	10.0
خبز	1.0	1.5	2.0	2.5	3.0	4.0	5.0
قهوة	6.0	8.0	10.0	30.0	50.0	90.0	110.0
سكر	3.0	4.0	6.0	9.0	12.0	30.0	40.0
صابون	3.0	4.0	12.0	10.0	15.0	30.0	40.0

المطلوب: 1- أوجد معدل النمو السنوي المتوسط لكل مادة.

2- أوجد الرقم القياسي ذو الأساس المتحرك و فسره.

قائمة المراجع

أولا ، باللغة العربية:

- 1- أنيس كنجو . الإحصاء وطرق تطبيقه في ميادين البحث العلمي . ج 1 . مؤسسة الرسالة . ط 2 . 1982 .
- 2- دومينيك سالفاتور . الإحصاء و الإقتصاد القياسي - سلسلة ملخصات شوم - دار ماكجر وهيل للنشر . د.م . الجامعية . 1983 .
- 4- د . عبدالعزيز هيكل . مبادئ الأساليب الإحصائية . دار النهضة العربية - بيروت . 1974 .
- 5- د . عبد القادر حلومي . مدخل الى الإحصاء . ديوان المطبوعات الجامعية - الجزائر . 1993 .
- 6- د . عصام عزيز شريف . مقدمة في القياس الإقتصادي . ديوان المطبوعات الجامعية - الجزائر . ط : 1981 .
- 7- محمد صبحي أبو صالح وعدنان محمد عوض . مقدمة في الإحصاء . ديوان المطبوعات الجامعية - الجزائر . 1984 .
- 8- المجموعة الإحصائية السنوية . د.و . للإحصائيات . 1994 .
- 9- ناظم حيدر - الوسيط في الإحصاء التطبيقي - دار الكتاب - 1977 ط 2 .
- 10- أحمد عبادة سرحان . العينات . القاهرة . بدون سنة .
- 11- عبدالسلام أبو قحف . التسويق مدخل تطبيقي . دار الجامعة الجديدة . الإسكندرية . 2002 .

ثانياً، باللغة الفرنسية:

- 1- **C.Labrousse.** Statistique Exercices Corrigés 1 . Dunod.1975.
- 2- **G.Calot.** Cours De statistique Descriptive .Dunod. Paris. 1975
- 3- **Hamdani Hocine.** Statistique Descriptive Et Expression Graphique. O.P.U.1998
- 4- **J.Lecaillon Et C.Labrousse.** Statistique Descriptive . Cujas. Paris.
- 5- **J.Leurion.** Statistique (Mathematique Appliquées). T2. Foucher. 1970.
- 6- **Murray R. Spiegel .** Theorie Et Applications De La Statistique.Serie Schaum . 1983.
- 7- **P. pacé Et R.Cluzel .** Statistique Et Probabilité . T1.Delagrave.
- 8- **Pierre Bailly.** Exercices corrigés de statistique descriptive. O.P.U. 1993

فهرس المواضيع

الموضوع	
مقدمة	1
الفصل الأول: مفاهيم، إستخدامات و منهجية	3
أولاً: مفهوم الإحصاء	3
1- تعداد	3
2- إحصائيات	3
3- علم الإحصاء	7
ثانياً: مجالات إستخدام الإحصاء	9
ثالثاً: أنواع البحوث الإحصائية	9
1- البحوث الإحصائية الوصفية	9
2- البحوث الإحصائية التحليلية	9
3- البحوث الإحصائية التجريبية	10
رابعاً: منهجية البحث الإحصائي	10
1- المرحلة الأولى: التحديد الدقيق للظاهرة المدروسة	10
2- المرحلة الثانية: جمع البيانات الإحصائية	13
أ- مصادر جمع البيانات	13
ب- أساليب جمع البيانات من المصادر المباشرة	14
ج- طرق جمع البيانات الإحصائية	21
د- الإستمارة الإحصائية	23
هـ- أخطاء جمع البيانات الإحصائية	25
و- مراجعة البيانات	25
3- المرحلة الثالثة: تبويب و عرض البيانات	27
4- المرحلة الرابعة: تحليل البيانات و إستقراء النتائج	27
أسئلة و تمارين	29
الفصل الثاني: تبويب البيانات الإحصائية	31
أولاً: مواصفات الجداول الإحصائية	31
ثانياً: الجداول التكرارية ذات الصفات النوعية	33
1- الجداول التكرارية البسيطة	33
2- الجداول التكرارية المزدوجة للصفات النوعية	35
ثالثاً: الجداول التكرارية ذات الصفات الكمية	36
1- الجداول التكرارية غير المستمرة	37
2- الجداول التكرارية الكمية المستمرة	39
رابعاً: أنواع التوزيعات التكرارية المستمرة	44
1- التوزيع التكراري المغلق	44
2- التوزيع التكراري المفتوح	44
3- التوزيعات التكرارية المتجمعة	45
4- التوزيع التكراري النسبي و المائوي	48
تمارين	51

55	الفصل الثالث: العرض البياني
55	أولاً: مواصفات الأشكال البيانية
56	ثانياً: الرسوم التكرارية
56	1- الأعمدة التكرارية
58	2- المدرج التكراري
63	3- المضلع التكراري
65	4- المنحنى التكراري
66	5- المدرج التكراري المتجمع
69	6- المنحنى التكراري المتجمع
70	ثالثاً: المنحنيات الزمنية
71	رابعاً: الشكل الدائري
73	خامساً: الشكل المستطيل
75	سادساً: الشكل القطبي
76	سابعاً: الطريقة التصويرية
77	تمارين
81	الفصل الرابع: مقاييس النزعة المركزية
82	أولاً: الوسط الحسابي
82	1- الوسط الحسابي للبيانات غير المبوبة
83	2- الوسط الحسابي للبيانات المبوبة
83	أ- البيانات المبوبة التي مدى فئاتها معدوم
84	ب- البيانات المبوبة التي مدى فئاتها أكبر من الصفر
87	3- خواص الوسط الحسابي
107	ثانياً: الوسيط
107	1- وسيط البيانات غير المبوبة
108	2- وسيط البيانات المبوبة
115	3- خصائص الوسيط
115	ثالثاً: المنوال
115	1- البيانات غير المبوبة و البيانات التي طول فئاتها معدوم
115	2- البيانات المبوبة التي مدى فئاتها أكبر من الصفر
118	3- خصائص المنوال
119	4- العلاقة بين المنوال و الوسيط و الوسط الحسابي
119	رابعاً: الوسط الهندسي
119	1- البيانات غير المبوبة
121	2- البيانات المبوبة
123	خامساً: الوسط التريعي
123	1- للبيانات غير المبوبة
124	2- للبيانات المبوبة
125	سادساً: الوسط التوافقي
125	1- للبيانات غير المبوبة
125	2- للبيانات المبوبة

126	3- العلاقة بين الوسط التوافقي والوسطين الهندسي والحسابي
126	سابعا: الربيعيات
127	1- الربيع الأدنى
127	2- الربيع الأعلى
128	ثامنا: العشريات
128	تاسعا: المؤينات
130	عاشرا: العلاقة بين الوسط والربيعيات والعشريات والمؤينات
133	تمارين
137	الفصل الخامس: مقاييس التشتت
138	أولا: مدى التغير
139	ثانيا: الانحراف المتوسط
139	1- الانحراف المتوسط للبيانات غير المبوبة
140	2- الانحراف المعياري للبيانات المبوبة
142	ثالثا: التباين، الانحراف، المعياري، العزوم
142	1- التباين
142	أ- للبيانات غير المبوبة ب- للبيانات المبوبة
144	2- الانحراف المعياري
147	3- العزوم
148	رابعا: الانحراف الربيعي
148	خامسا: معاملات التشتت النسبية
148	1- معامل الاختلاف
149	2- الانحراف الربيعي النسبي
149	سادسا: العلاقة بين بعض مقاييس التشتت النسبية
150	تمارين
153	الفصل السادس: أشكال التوزيعات التكرارية
153	أولا: التماثل التام
155	ثانيا: الالتواء
156	1- قيمة الالتواء
157	2- معامل بيرسون للالتواء
162	3- معامل الالتواء الربيعي
162	4- معامل الالتواء العزمي
162	ثالثا: التفرطح
163	1- معامل فيشر
164	2- معامل كيللي
166	رابعا: أشكال أخرى
166	1- المنحنى الرائي
166	2- المنحنى اللامي
167	3- المنحنى متعدد القمم
167	4- منحنيات القطع المكافئ
168	أسئلة و تمارين

169	الفصل السابع: الإنحدار والارتباط
171	أولا: الإنحدار
171	1- الإنحدار الخطي البسيط
187	2- الإنحدار الثلاثي
190	ثانيا: معاملات الارتباط
190	1- معامل الارتباط الخطي البسيط
193	2- معامل الارتباط الجزئي
195	3- معامل ارتباط الرتب
197	تمارين
203	الفصل الثامن: السلاسل الزمنية
204	أولا: أشكال تغيرات السلسلة
204	1- التغيرات طويلة المدى 2 - التغيرات الموسمية
205	3- التغيرات الدورية
205	4- التغيرات العشوائية
205	ثانيا: الحشونة
207	ثالثا: تقدير الاتجاه العام للسلسلة
207	1- الشكل الخطي و شبه الخطي
219	2- الشكل غير الخطي للاتجاه العام
228	رابعا: تحديد المركبات الموسمية و حذف تغيراتها
228	1- طريقة النسبة الى الاتجاه العام
231	2- طريقة متوسطات الموسم
237	تمارين
241	الفصل التاسع : الأرقام القياسية و معدلات النمو
241	أولا: الأرقام القياسية
241	1- الأرقام القياسية البسيطة
242	أ- طريقة المناسيب البسيطة ب- طريقة الوسط الحسابي للمناسيب البسيطة
243	ج- طريقة الوسط الهندسي للمناسيب البسيطة
244	د- الطريقة التجميعية البسيطة
245	2- الأرقام القياسية المرجحة
246	أ- الطريقة التجميعية المرجحة
248	ب- طريقة لاسير ج- طريقة باش د- طريقة فيشر
250	3- الأرقام القياسية ذات الأساس المتحرك
251	4- إختيار فترة الأساس
252	5- خصائص الأرقام القياسية الموضوعية
254	ثانيا: معدلات النمو
255	1- معدل النمو البسيط
255	2 - معدل النمو السنوي المتوسط
257	تمارين
261	قائمة المراجع
263	فهرس المواضيع